Vaasan yliopisto
UNIVERSITY OF VAASA

Mimmi Huusko

Defining the operational needs and optimal process coverage of digital twin in ABB Drive Products' operations

School of Technology and Innovations
Master's thesis in Industrial Engineering and Management
Industrial Systems Analytics

Vaasa 2025

UNIVERSITY OF VAASA
School of Technology and Innovations
Author: Mimmi Huusko
Title of the Thesis: Defining the operational needs and optimal process coverage of digital twin in ABB Drive Products' operations
Degree: Master of Science in Technology
Programme: Industrial Systems Analytics
Supervisor: Rayko Toshev **Instructor:** MSc. Tech Timo Rissanen
Year: 2025 **Pages:** 102

ABSTRACT:
Digital twin technology is becoming an increasingly recognized tool in manufacturing optimization. It is a vital component of an Industry 4.0 strategy, integrating advanced technologies such as artificial intelligence, cloud computing, and the Industrial Internet of Things to provide enhanced process visibility, improved operational efficiency, and scenario simulation capabilities.

Digital twins are designed for a specific environment and purpose. Consequently, the operational needs must be defined in advance to create an optimal solution recommendation and assess the possible benefits of the system before deciding on implementation. To evaluate the applicability and potential of digital twin technology within the case company's operations, this study conducted semi-structured interviews with seven stakeholder groups. The interviews focused on identifying current operational challenges, data requirements, and expected system functionalities. The results were analyzed by identifying key words in thematic categories.

The research identifies that to address current operational challenges, the system must provide real-time visibility into production processes and material flow, coupled with planning, forecasting, decision support, analytics, and process optimization capabilities. To meet these needs effectively, a predictive and prescriptive digital twin model is recommended, with potential enhancement through real-time control features. Additionally, the study determines that an optimal digital twin should be modular, scalable, and tailored to varying user needs, focusing primarily on high-level process flow and workstation status enhanced with more granular disruption details and material flow data. Product and component-level digital twins are considered secondary in priority for current operational improvements.

For optimizing the case company's manufacturing operations, two levels of analytics are proposed: predictive and prescriptive. Predictive capabilities support disruption management and proactive problem-solving, while a prescriptive twin offers strategic decision-making support and process optimization. However, findings reveal that a digital shadow – a solution that collects, analyzes, and visualizes data without direct feedback to the physical system – suffices to address the identified operational needs. While a full digital twin could offer additional benefits, particularly in automation, the current operational environment would significantly benefit from implementing a digital shadow.

KEYWORDS: Digital twin, manufacturing optimization, Industry 4.0, predictive analytics, prescriptive analytics, process optimization, operational needs

VAASAN YLIOPISTO
Tekniikan ja innovaatiojohtamisen yksikkö
Tekijä: Mimmi Huusko
Tutkielman nimi: Defining the operational needs and optimal process coverage of digital twin in ABB Drive Products' operations
Tutkinto: Diplomi-insinööri
Oppiaine: Industrial Systems Analytics
Valvoja: Rayko Toshev **Ohjaaja:** MSc. Tech Timo Rissanen
Valmistumisvuosi: 2025 **Sivumäärä:** 102

TIIVISTELMÄ:
Digitaalisesta kaksosesta on tullut yhä tunnetumpi työkalu tuotannon optimoinnissa. Se on merkittävä osa Industry 4.0 -strategiaa, yhdistäen edistyneitä teknologioita kuten tekoälyllä toteutettavaa analytiikkaa, pilvilaskentaa, ja teollista esineiden internetiä (IIoT), tarjoten parannettua prosessin läpinäkyvyyttä, operatiivista tehokkuutta ja skenaarioiden simulointia.

Digitaaliset kaksoset suunnitellaan aina tiettyyn ympäristöön ja käyttötarkoitukseen. Sen vuoksi on ensin määriteltävä operatiiviset tarpeet, jotta voidaan luoda optimaalinen teknologinen ratkaisu ja arvioida järjestelmän mahdollisia hyötyjä ennen kuin tehdään päätös toteutuksesta. Tämän tutkimuksen tavoitteena oli arvioida digitaalisen kaksosen soveltuvuutta ja potentiaalia kohdeyrityksen toiminnassa. Tutkimusta varten suoritettiin puolistrukturoidut haastattelut seitsemän sidosryhmän edustajille. Haastatteluissa kartoitettiin tämänhetkisiä operatiivisia haasteita, datavaatimuksia, ja järjestelmältä odotettuja hyötyjä. Tulokset analysoitiin etsimällä avainsanoja temaattisista kategorioista.

Tutkimus osoittaa, että nykyisten operatiivisten haasteiden ratkaisemiseksi järjestelmän tulee tarjota reaaliaikainen näkyvyys tuotantoprosesseihin ja materiaalivirtaan, yhdistettynä suunnittelu-, ja ennustuskyvykkyyteen, päätöksenteon tukemisen kyvykkyyteen, analytiikkaan ja prosessien optimointikyvykkyyksiin. Näiden tarpeiden täyttämiseksi suositellaan ennakoivan ja määräävän (prescriptive) digitaalisen kaksosen mallia, mahdollisesti täydennettynä reaaliaikaisella ohjausominaisuudella. Lisäksi tutkimus osoittaa, että optimaalisen digitaalisen kaksosen tulisi olla modulaarinen, skaalautuva ja käyttäjätarpeisiin mukautuva, ja painottaa erityisesti korkean tason prosessivirtatietoja ja työpisteiden statustietoja, joita täydentävät tarkemmat häiriöiden ja materiaalivirran tiedot. Tuote- ja komponenttitason digitaaliset kaksoset katsotaan toissijaisiksi nykyisten operatiivisten parannusten näkökulmasta.

Kohdeyrityksen tuotannon optimointia varten esitetään kahden tason analytiikkaa: ennakoivaa (predictive) ja määräävää (prescriptive) analytiikkaa. Ennakoivat kyvykkyydet tukevat häiriöiden hallintaa ja ennakoivaa ongelmanratkaisua, kun taas määräävä digitaalinen kaksonen tarjoaa strategisen päätöksenteon tukea ja prosessien optimointia. Tutkimuksen löydökset kuitenkin osoittavat, että digitaalinen varjo – ratkaisu, joka kerää, analysoi ja visualisoi dataa ilman suoraa palautetta fyysiseen järjestelmään – riittää tunnistettujen operatiivisten tarpeiden täyttämiseen. Vaikka täysimittainen digitaalinen kaksonen voisi tarjota lisähyötyjä erityisesti automaation yhteydessä, nykyinen operatiivinen ympäristö hyötyisi merkittävästi digitaalisen varjon käyttöönotosta.

AVAINSANAT: Digitaalinen kaksonen, tuotannon optimointi, teollisuus 4.0, ennakoiva analytiikka, määräävä analytiikka, prosessin optimointi, operatiiviset tarpeet

Contents

Pictures

Figures

Tables

Abbreviations

AGV	Autonomous Guided Vehicle
AI	Artificial Intelligence
AMR	Autonomous Mobile Robot
SCADA	Supervisory Control and Data Acquisition
CPS	Cyber-Physical System
DL	Deep Learning
Drive	Variable-speed drive
ERP	Enterprise Resource Planning system
FEL	Future Event List
ICS	Industrial Control Systems
IoT	Internet of Things
IIoT	Industrial Internet of Things
MES	Manufacturing Execution System
ML	Machine Learning
OEE	Overall Equipment Effectiveness
OPF	One Piece Flow
RFID	Radiofrequency identification
SAS	Symbiotic Autonomous System
UWB	Ultra-Wide Band
VSM	Value Stream Mapping
XR	Extended Reality

1 Introduction

Digital twin technology has become increasingly recognized as a revolutionary tool in the manufacturing field, offering the potential for improved process control, predictive maintenance, and data-driven decision-making. This chapter introduces the background and context of the study, defines the research problem and objectives, and presents the scope and significance of the research. The structure of the thesis is also summarized.

1.1 Background and context of the study

The technological advancements of the 21st century have encouraged manufacturing plants to transform their operations towards Smart Factories, and one of the key enabling technologies is a digital twin. Digital twin is a virtual representation of a physical product, process, or a system, and it provides a platform for real-time monitoring, simulation, and optimization. It has significant potential for enhancing overall operational efficiency across the system lifecycle. An effective digital twin solution must address specific challenges that are unique in every manufacturing environment and for every stakeholder. Despite the promising benefits, a successful implementation requires a clear understanding of operational needs. There is no such thing as a general digital twin; every solution is designed for a specific purpose. Without understanding the current pain points and expectations, organizations risk investing in a solution that does not deliver the expected benefits or fails to align with business goals.

Foundational literature by Grieves (2016) identifies three core elements: the physical entity, the digital replica, and the data connecting them. In this study, the digital twin is framed as a socio-technical system embedded within the operational context of the case company. Relevant theories include Cyber-Physical Systems (CPS), which provides a systems-level understanding of how physical processes and digital technologies interact (Boyes et al., 2018), and Systems theory, which highlights the interdependence between processes and the need for comprehensive modeling in operations (J. Liu et al., 2021).

Understanding the case company's operational needs requires identifying the current process inefficiencies and data limitations. This research draws on the technology-organization-environment (TOE) framework to explore the drivers and constraints of adopting a digital twin. This ensures that multiple design aspects are considered before implementation (Nguyen et al., 2022). Related key constructs include process visibility and control, decision support, and integration requirements.

The concept of process granularity refers to the level of detail captured in digital representations. A similar concept is fidelity, which describes the accuracy and completeness of the representation. In this research, the two terms are used interchangeably to ensure consistency and avoid loss of information. Coverage refers to the scope of the modelled processes. This study utilizes the Fit-For-Purpose approach to determine the appropriate depth of the solution. Another related theory is Business Process Management (BPM), which guides how processes should be identified, modeled, and improved (Dumas et al., 2018). Additionally, Model Theory suggests that the level of detail in the model should match the defined needs without overcomplicating the system (Maier et al., 2017).

The research aims to investigate how digital twin capabilities can be leveraged to enhance operational performance. The study uses principles from Operations Research and Data-Driven Decision Making (DDDM) to explore optimization methods. Some supporting frameworks include Lean, which focuses on waste elimination and efficiency. Optimization domains include for example predictive maintenance and bottleneck analysis.

1.2 Research problem and objectives

ABB Drive Products, a key player among the variable frequency drive manufacturers, is considering implementing digital twin technology to optimize its operations. The aim of this research is to aid the case company in determining whether a digital twin would create advantages in their manufacturing operations. To accomplish that goal, the operational needs have to be visible and clearly defined. A systematic approach is needed to capture the operational needs and system expectations. The possible advantages of a digital twin are highly dependent on the particular twin solution, which is why the optimal scope, granularity, and level of automation are also researched. Without a well-defined framework, there is a risk of either excessive complexity or a system that is inadequate. Three research questions are created to guide this research in understanding the operational needs, optimal scope and granularity, and optimization potential of a digital twin:

- **RQ1**: What are the operational needs for a digital twin solution?
- **RQ2**: What is the optimal process coverage and granularity for the digital twin based on the operational needs?
- **RQ3**: How can digital twin be used in optimizing the ABB Drive Products operations?

The objective is to

- Provide a comprehensive understanding of digital twin technology and its capabilities,
- Understand the operational needs of the manufacturing environment,
- Determine the optimal process coverage and granularity,
- Assess the potential of optimizing the operations with a digital twin.

1.3 Scope and significance of the research

This research aims to define a framework that guides the company in deciding the implementation scope and application feasibility for a digital twin, without fully examining the technical development or developing a prototype. The framework culminates in a conceptual model linking operational needs (input), digital twin design parameters (process coverage & granularity), and optimization outcomes (output). A continuous feedback loop ensures that the insights gained from optimization outcomes inform on-going adjustments to operational needs and design parameters (Figure 1). The scope is limited to the evaluation of operational needs and how a digital twin can address these needs and deliver measurable improvements in operations' quality, efficiency, and adaptability.

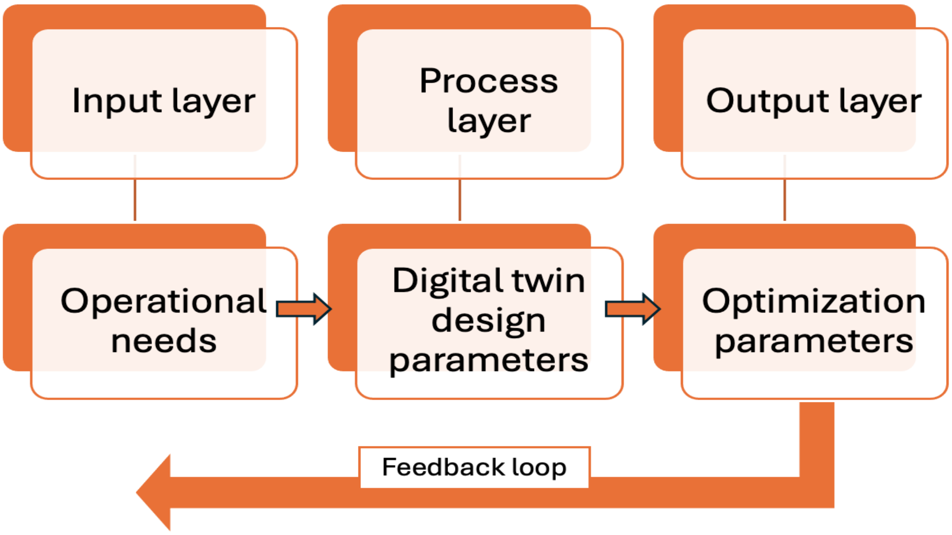


Figure 1. Digital twin implementation framework.

For ABB Drive Products, incorporating digital twin technology creates an opportunity for addressing key challenges in operational efficiency, product quality, and production flexibility. Understanding the operational requirements for a digital twin solution (RQ1) is crucial, because it ensures that a possible future digital twin development is aligned with business and production goals, rather than serving as a generic technological addition. Furthermore, defining the appropriate process coverage and granularity (RQ2) ensures that the company is able to balance the depth of required data for meaningful insights with the cost and complexity of system integration. Lastly, identifying potential use cases for digital twin technology in the manufacturing environment (RQ3) will highlight areas where measurable improvements can be seen, such as identifying and reducing bottlenecks, enhancing decision-making, and supporting lean manufacturing principles.

The insights learned from this research have implications not only for ABB Drive Products, but also for other manufacturing enterprises that consider adopting a digital twin or similar solutions.

1.4 Structure of the thesis

The rest of this thesis is structured as follows:

- Chapter 2: Literature review – Provides an overview of the manufacturing industry, key theoretical concepts of digital twin technology, and its applications in industrial systems.
- Chapter 3: Methodology – Details the research design, data collection methods, and analysis techniques. Reliability of the research is also discussed.
- Chapter 4: Analysis – Analyses the qualitative data from the focus group interviews in four thematic categories.
- Chapter 5: Results and Conclusions – Presents and interprets findings from the interviews, summarizes key findings and answers the research questions.
- Chapter 6: Discussion and Recommendations – Discusses potential benefits against literature and implications for the case company and presents suggestions for future research.

By following this structure, this thesis aims to create value for the case company as a strategic decision-making tool.

2 Literature review

This section covers the research related to digital twins in manufacturing. The purpose of this section is to provide an introduction to manufacturing operations, a comprehensive view of the history and evolution of digital twin, and how digital twins support the manufacturing industry. The case company's manufacturing operations are explained to gain context of this specific case and the applicability of digital twin in this specific environment. Implementation process and the prerequisites are described to underline the importance of adequate design and planning. Finally, this section introduces five case studies of digital twin implementation and identifies some companies that have implemented digital twin technology in their operations.

2.1 Industry overview

Traditionally, manufacturing operations have been designed to process raw materials into finished products. With automation and digital twins, these processes are transforming from primary processes to smart processes, and are able to become interactive with other elements and their surroundings (M. Liu et al., 2021). Manufacturing operations can be illustrated as eight consequent steps, which are

1. Design,
2. Material procurement,
3. Production planning,
4. Manufacturing,
5. Quality control,
6. Assembly,
7. Testing and inspection, and
8. Delivery.

These steps are interconnected and interdependent, creating a complex manufacturing environment (Gupta & Starr, 2014). Some other characteristics that complex systems have are emergent and non-linear behavior, positive and negative feedback loops, sensitivity to initial conditions, adaptation, and self-organization (Yuan et al., 2024). Traditional process optimization and quality management tools like Lean, Six sigma, and Agile are as relevant, but the 21st century manufacturing environments create the need to utilize proper systems modelling tools to understand, supervise, and manage the complex environment. Lean manufacturing is a concept that utilizes tools like Value Stream Mapping (VSM), continuous improvement, and Just In Time (JIT) production (Gupta & Starr, 2014). In agile production, the process is designed to accommodate changing requirements. Work is delivered in small, functional increments to ensure flexibility and responsiveness (Gupta & Starr, 2014). Six Sigma is a quality management approach that includes tools such as DMAIC, root cause analysis, statistical process control, and Failure Modes and Effects Analysis (FMEA) to improve process performance and reduce defects (Skorupińska et al., 2024).

The fourth industrial revolution, or Industry 4.0, is the next stage in the development of industrial best practices. Following Industry 3.0 which introduced robotics and Enterprise Resource Planning (ERP) systems, the new stage includes technologies such as Cyber Physical Systems (CPS), Industrial Control Systems (ICS), Supervisory Control and Data Acquisition (SCADA), and Industrial Internet of Things (IIoT) (Boyes et al., 2018). The five key technologies at the basis of digital twin implementation are IoT, CPS, Cloud computing (CC), Edge computing, Big data, and AI (Lattanzi et al., 2021).

2.1.1 ABB Drive Products

ABB Drive Products can be divided into three production areas - Small & Medium drives, Late configuration drives, and Large drives. They vary in process times, inbound and outbound logistics methods, and levels of automation, but they all utilize the same one-piece flow (OPF) production model with 4-6 synchronized workstations. In addition to OPF lines, the production area includes supporting operations like pre-manufacturing stations (kitting cells) and automation cells. The manufacturing starts when an operator selects the next drive from the work queue and prints a rating plate. Then the materials are picked from the shelf or a wooden pallet, traced to the Manufacturing Execution System (MES), and assembled according to work instructions, using smart screwdrivers. When the first operation is completed, the operator checks the operation as complete and moves the drive to the next workstation to be processed. In general, the last workstation moves the drive to a testing pallet, connects testing cables to the drive and calls for an Autonomous Mobile Robot (AMR) which carries the drive to the testing system. The testing systems differ between production areas: Late configuration production line and LD2 production area have manual testing process without automation, while other areas have automated manipulators and testing queues. After a test has been passed, the testing system transfers the drive to a packing area, where final assembly and packing is completed.

2.2 Foundation and conceptual background of digital twin

Simulation for manufacturing systems was first invented in the 1960s when NASA created digital models of space systems, one of the most popular modelled systems being Apollo 13. This technology became more popular in the 1970s when the technology expansion period first started. Real-time simulation has been possible from the 1990s, and the first mention of a “digital twin” was in 1997 by Hernández and Hernández (Mourtzis, 2020; M. Liu et al., 2021). Starting from the early 2000s, the digitalized era has elevated digital twins from being offline simulations to allowing two-way communication between the physical and digital world, and integrating artificial intelligence, enabling forecasting and optimization suggestions (Mourtzis, 2020).

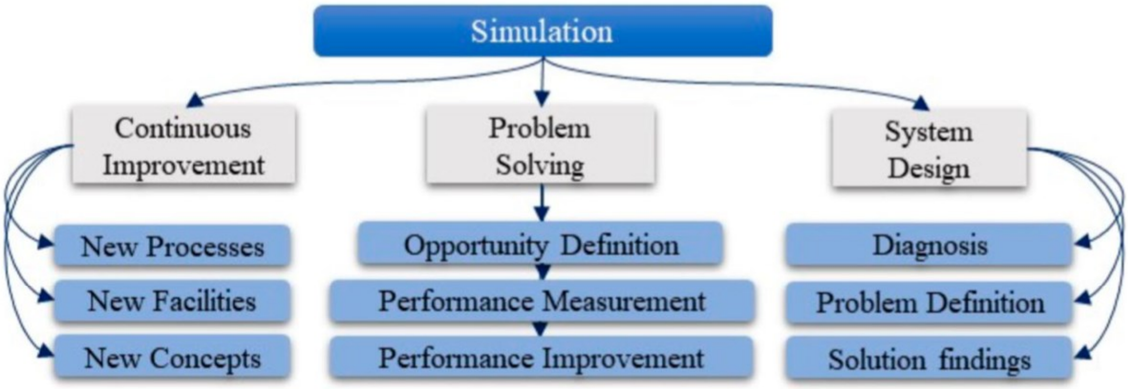


Figure 2. Benefits of simulation in the design and operation of manufacturing systems (Mourtzis, 2020).

The main added value from simulation in the design and operation of manufacturing systems can be categorized into continuous improvement, problem solving, and system design capabilities (Figure 2). Continuous improvement refers to the new process, facility, and concept development, all of which achieve significant benefits when modelled virtually before implementing them in the physical world. Opportunity definition, as well as performance measurement and improvement are also capabilities that simulation introduces to the operation of manufacturing systems. In the system design phase, it is possible to make diagnoses, define problems, and find solutions via simulation, which reduces the trial-and-error period of system development. (Mourtzis, 2020)

Mourtzis (2020) creates a comprehensive view of simulation possibilities in product and production lifecycle by listing 12 different simulation tools (Figure 3). These tools include computer-aided design, -manufacturing, and -process planning (CAD, CAM, & CAPP), digital mock-ups, material flow-, process-, ergonomics-, supply chain-, and layout planning simulations, manufacturing execution systems (MES), supervisory control and data acquisition (SCADA), and design and planning of manufacturing networks. Traditionally these tools have been used separately and for one purpose, but with 4th industrial revolution advances, it is possible to simulate complex systems and manufacturing plants as a whole by merging these tools together in one system and creating a digital twin (Mourtzis, 2020).

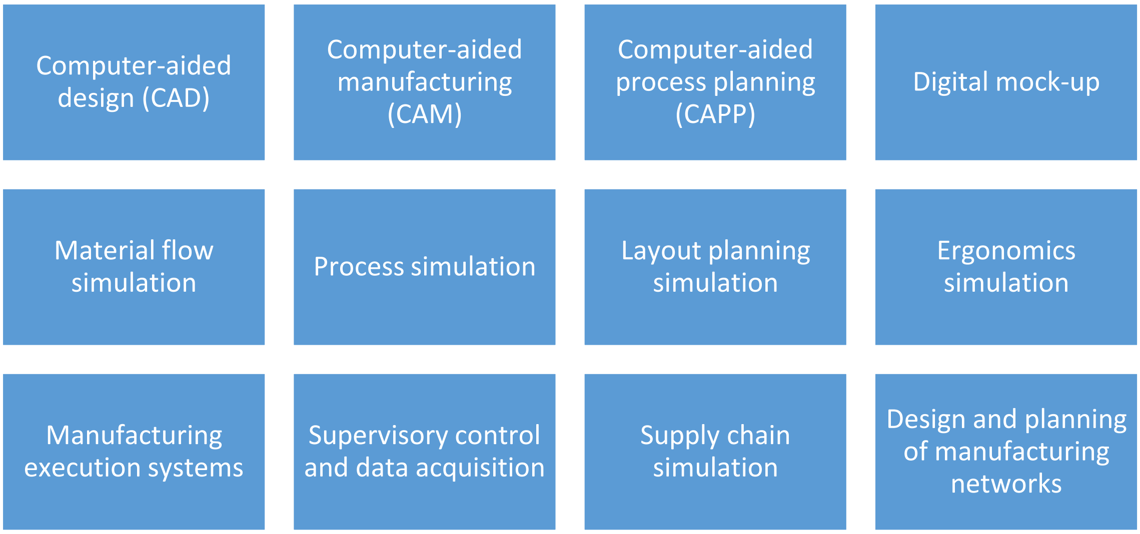


Figure 3. Simulation tools in product and production lifecycle (Mourtzis, 2020).

The purpose of simulation-based optimization is finding the optimal inputs without evaluating every possibility. Discrete-event simulation is common and suitable for high-fidelity manufacturing flow modelling, where data is derived from IIoT or operational systems which manage business or manufacturing data (ERP and MES). Shifting bottlenecks, shared resources, and complex scheduling rules make it difficult to create a low-fidelity simulation model, while high-fidelity models require massive computational resources and take more time. Multi-fidelity models ensure that low-fidelity models first create the earliest feasible solution which then aid in the determining of optimal solution. Simulation-based optimization models need real-time data, which is why it is commonly implemented after a virtual replica or even a digital twin has been created. (Zhang et al., 2022)

In 2002 digital twin became a commonly researched topic in the manufacturing industry following the works of Michael Grieves (Grieves, 2016). A general definition by VanDerHorn & Mahadevan (2021) states that a digital twin can be described as *“a virtual representation of a physical system (and its associated environment and processes) that is updated through the exchange of information between the physical and virtual systems.”* With the recent Industry 4.0 advancements in manufacturing and data management, it is now becoming one of the most significant technologies for manufacturers. Simultaneously, the products and processes of manufacturing have become more complex, and digital twin needs to adjust accordingly. Artificial intelligence (AI) combined with the modelling and simulation of a “traditional” digital twin enables the intelligent digital twin to not only replicate the physical universe, but also to have goal-seeking, active behaviour (Grieves, 2022). In practice, this enables digital twin to assist human behaviour by creating recommendation for next actions based on the goals that have been defined. The assistance is based on the two superior qualities that computers have against humas: the ability to perform complex calculations and ruling out human bias (Grieves, 2022).

Digital twins consist of three elements: physical element, virtual counterpart, and the communication channel between them (Figure 4). According to Attaran & Celik (2023), there are four core technologies that create the basis of a digital twin – the Internet of Things (IoT), Cloud Computing, Artificial Intelligence (AI), and Extended Reality (XR). These technologies ensure real-time data collection and storage, create insights from information, and form a digital representation of an object. IoT is a primary technology in all digital twin applications. It refers to the network of connected devices and the supporting technology. IoT collects real time data with sensors. Cloud computing allows for the data to be stored and an easy access for the stored data over the internet. It reduces computation time and provides an efficient way to gather large amounts of data. AI utilizes machine learning and provides automatic data analysis, outcome prediction, and suggestions. XR includes augmented reality, virtual reality, and mixed reality, which merge the virtual and physical world, providing a platform where digital and physical objects can co-exist and interact in real time. (Attaran & Celik, 2023)

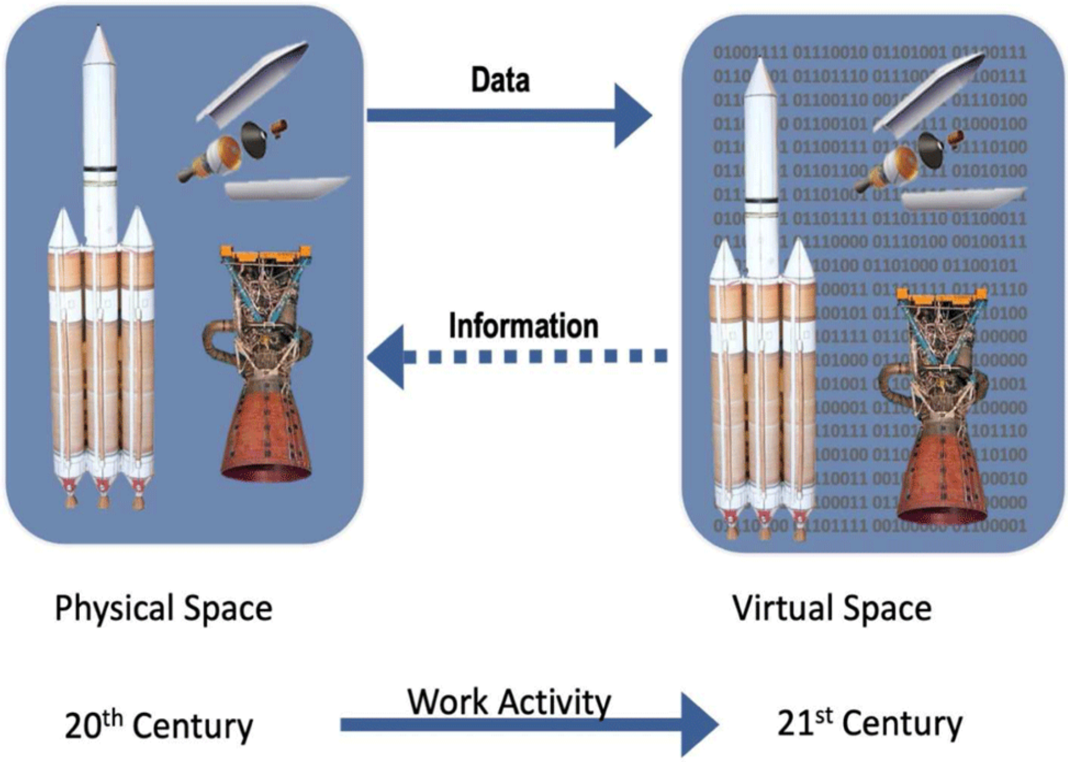


Figure 4. Digital twin model (Grieves, 2022).

Digital twin is associated to all four parts of a product lifecycle, from creation and building to operating, and finally to disposal phase. It is considered to be the most valuable during the creation of new products, systems and processes (Grieves, 2022). Grieves (2022) divides digital twins to three types depending on the use case: prototype (DTP), instant (DTI) and aggregate (DTA). DTP is a virtual prototype of a product that does not yet exist, and it can potentially minimize development time and resource waste, for example energy and material use (Grieves, 2022). The impact of design and creation phase on future product cost is approximated to 80% (C. Lehner et al., 2024), which highlights the value of a prototype twin. DTI is a twin to the physical asset and is linked to it throughout the whole life cycle. It is used to test different usage scenarios and optimize the asset's operation. DTA gathers data from multiple DTIs and simulates group behavior to understand overall functionality. Digital thread is a similar concept to digital twin, but it includes the whole system or product lifecycle. It facilitates data stream connections and creates a comprehensive view of all related functional parts (Daase et al., 2023). The purpose is traceability and the seamless information access and sharing between users and organizations (C. Lehner et al., 2024).

2.2.1 Different types and granularities of digital twin

Digital twins can be created at varying levels of data integration, automation, and fidelity, and they are built with different scopes depending on the specific needs of a system that is being modeled. Digital twin classification by IoT Analytics (Hasan, 2023) concluded that there are at least 210 different combinations of qualities for a digital twin solution: 6 life-cycle phases x 5 hierarchical levels x 7 most common purposes (Figure 5). The modeled life-cycle phases are design, build, operation, maintaining, optimizing, and decommissioning. Possible hierarchical levels include information, component, product, system, and multi-system, which aims to describe the required level of detail. Hasan (2023) lists the most common purposes as digitize, visualize, simulate, emulate, extract, orchestrate, and predict. However, it is common that more than one combination is used in a twin design. The research also found that across 100 case studies the six most common applications were system prediction, system simulation, asset interoperability, maintenance, system visualization, and product simulation.

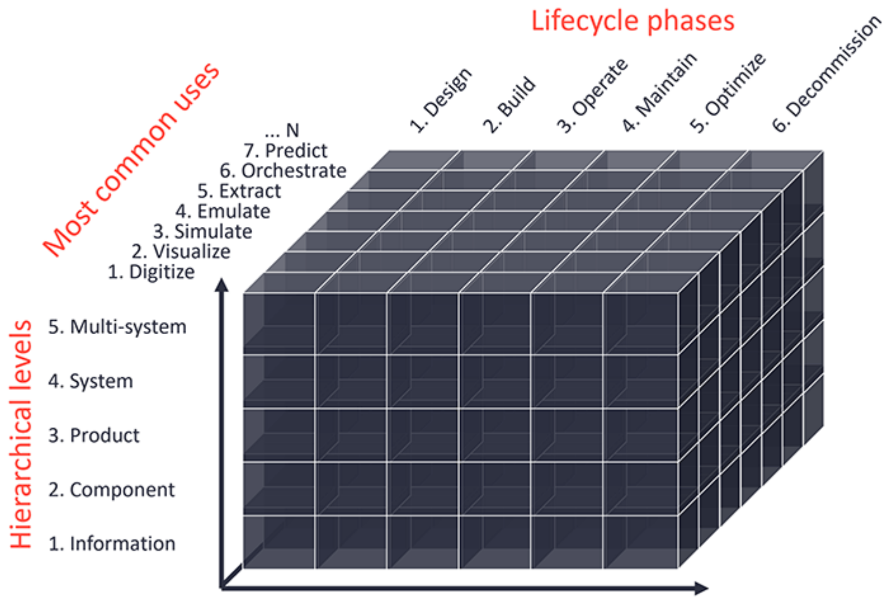


Figure 5. 210 different combinations of digital twin characteristics (Hasan, 2023).

The research by Xie & Wan (2023) demonstrates that detailed modeling offers a high degree of accuracy and can create substantial impact on the effectiveness of a digital twin in manufacturing process optimization. On the other hand, as Kober, Algan, et al. (2023) conclude, high accuracy can introduce complexity, be more time-consuming in design and in operation, and require substantial computational resources. Some design characteristics that require careful consideration are

(1) Data integration: model, shadow, or twin
(2) Fidelity: no of parameters, accuracy, and abstraction
(3) Level of analysis: descriptive, diagnostic, predictive, prescriptive, or autonomous
(4) Scope: system, process, product, or component
(5) Extent of automation: the role allocation between human and machine.

Digital twins vary in data integration and can be divided into three different categories based on data integration: digital model, digital shadow, and digital twin. In Figure 6 they are represented as arrows, because in addition to being separate categories, they also build upon each other. Pronost et al. (2023) divide the digital twin in four subcategories, adding pre-digital twin which represents a twin without the physical counterpart. In their research they note that pre-digital twin is mainly used in iterative optimization, virtual evaluation and verification, process evaluation and optimization, which do not necessarily need a physical counterpart. A digital model is a virtual counterpart of a physical system, which has no automatic data flow to either direction, for example, a CAD layout. Digital shadows are the most popular category, and they are often falsely labeled as digital twins (Resman et al., 2021). They have automatic data flow from the physical system to its virtual counterpart, creating a real-time view of the modelled system. However, in digital shadow the data flow from virtual to physical system is not automatic but is done manually.

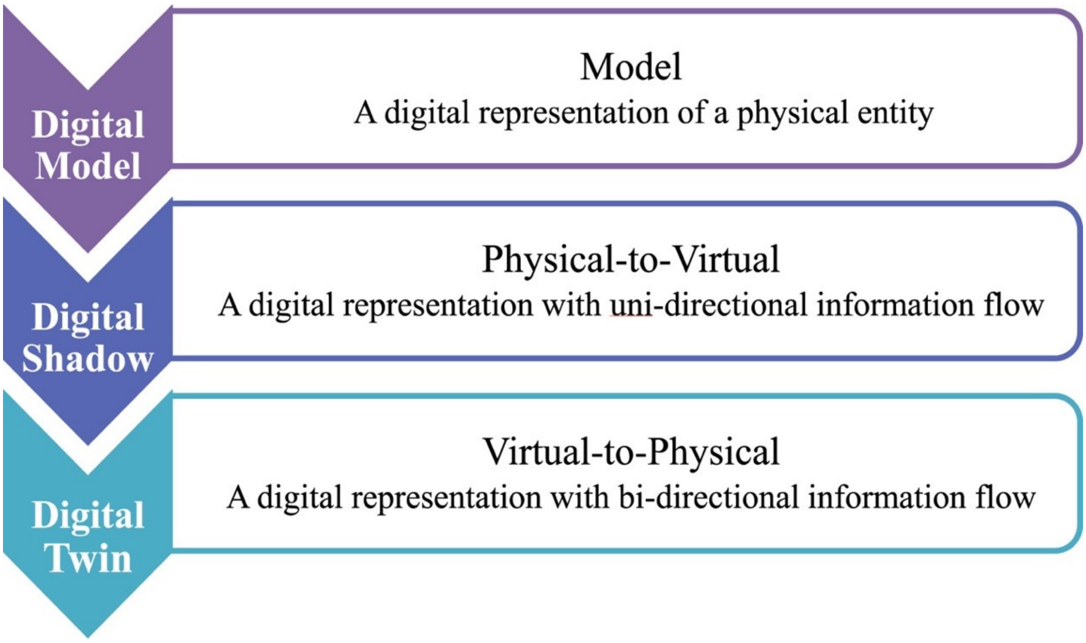


Figure 6. Three levels of data integration (Attaran & Celik, 2023).

Digital twin is the only category that is defined by a two-way, automated data flow (Resman et al., 2021). Digital twins are created by capturing or initiating a physical model and creating a digital model or a shadow, but instead of a one-way data flow from physical to virtual, they transfer data in both directions (Pronost et al., 2023). In practice, this means that the digital twin can change physical world parameters and for example adjust manufacturing line speed if it detects that it is not running optimally. Generally, a digital shadow is sufficient in the manufacturing industry, and it is the most commonly used version of "twins" (Pronost et al., 2023). They utilize real-time data, are able to analyze historical data, and can be enhanced with AI and other optimization tools. They are also capable of notifying users in cases of disruption, but without a bilateral data flow, cannot be used in manual or automated control of the physical system (Pronost et al., 2023).

Another, more extensive description of digital twin and especially simulation capabilities is fidelity. It describes the level of detail in a system, i.e. How accurately the digital twin models the physical world (Kober, Fette, et al., 2023). Generally, elevated fidelity results in higher costs, but to reach the goals that have been set for the digital twin, a high-fidelity model might be appropriate. Currently, the research is focused on multi-fidelity modeling, which uses simulation data from less accurate low-fidelity models, and then utilizes high-fidelity data for the improvement of recognized problem areas (Conti et al., 2023).

Level of analysis refers to the classification into descriptive, diagnostic, predictive, prescriptive, and autonomous system. It is closely linked to fidelity, since the amount and quality of data creates constraints for the analysis. Figure 7 describes four levels of analytics, and their value and difficulty compared to each other. A descriptive twin collects and presents real-time and historical data. The system can identify patterns, but usually they are reported on a dashboard and the interpretation is left for the user. The purpose of this analytics is to describe what has happened. A diagnostic twin utilizes data analytics to create trends and identify cause and effect relationships. The goal is to identify why something has happened, so that users are able to generate adequate solutions. With a predictive twin, the data analytics is used to predict future issues based on historical data. This is possible by utilizing AI and complex algorithms, and it creates opportunities for example for predictive maintenance. The prescriptive twin is able to not only diagnose and predict issues but also notify the user with alarms and create suggestions of future actions. These recommendations can include advice on how to prevent a future event, how to adjust to future event, or how to navigate current disruption. (Lorenz et al., 2024)

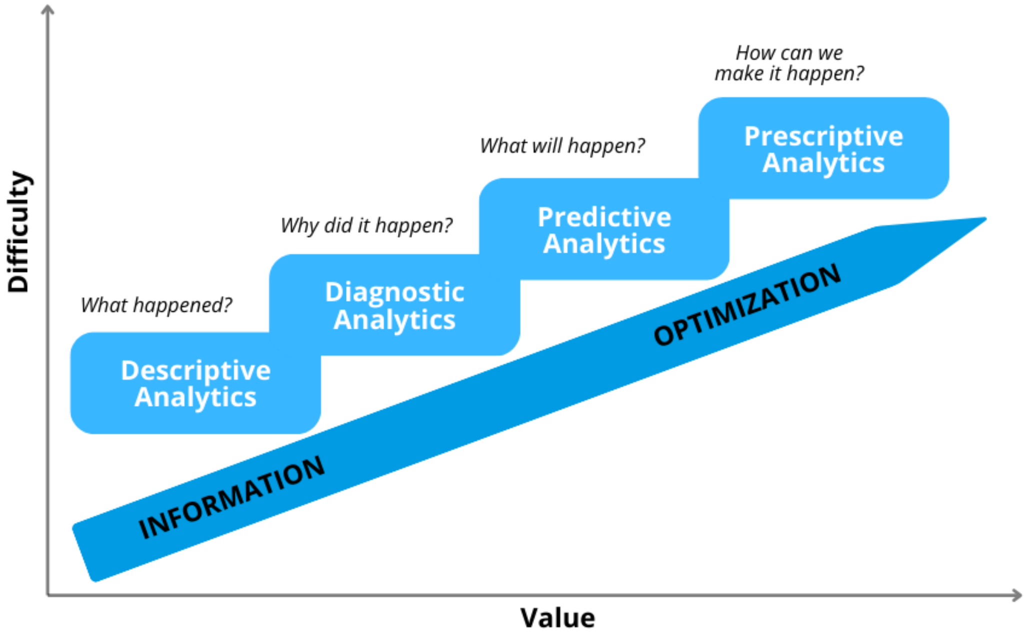


Figure 7. Analytics ladder.

Autonomous twin analyses the information, decides on how to proceed, and is able to implement the decision to the physical world (Xia et al., 2025). For example, an autonomous twin could adjust the temperature of the factory if it sees fit. Generally, descriptive and diagnostic twins are less complex and can be labeled as low-fidelity solutions, while a high-fidelity twin includes multiple data sources, which is high maintenance. For example, a descriptive analytics model is utilized for production flow to determine the current bottlenecks of a production system, and they are recognized as automatic pressing cell when product type X is manufactured, and the last work center of one production line when product type Y is manufactured. Then, prescriptive analytics is utilized for the determination of optimal manufacturing schedule of product types X and Y. This way, the whole production system does not need to be modeled with extreme fidelity. Additionally, the scope of the modeled system affects the complexity and the optimal level of fidelity, and is important to define based on the business and stakeholder needs. Anaya et al. (2024) divide the different possible scopes of digital twin in four different categories based on the purpose of the twin: process level, system level, asset level, and component level.

The extent of automation is another characteristic that needs to be defined. Modern digital twins are able to handle (collect, process, represent) data, analyze it against historical data, make simulations about every possible scenario, and even create recommendations on future actions and conduct them (Agrawal et al., 2023). This ability does not suggest that digital twins should always operate in that extent to be successful. Unfit role allocation between humans and digital twins can result in reduced monitoring, disadvantageous cost-benefit tradeoffs, and over-reliance on digital twins (Agrawal et al., 2023). Even if fully autonomous twin was the goal, realistically all current and foreseeable digital twin systems rely on human assistance, which is why it is important to define the optimal roles of automation and human work.

Example: Product variation has grown in a one-piece flow manufacturing line. A process engineer wants to check if the operation time in each work station could be more balanced. They need to collect, process, and represent the data, then evaluate if the process times are optimal, and then decide whether the balancing should be according to the top 75% of volume, or if they should take variation more into account and balance according to the average product processing time. Finally, they need to adjust the balance by updating work instructions and material allocation, moving materials, and updating possible quality documents. There are multiple possible automation scenarios. Some of the tasks could use digital twin to help with data analysis, simulation, or even executing tasks – either partially automating roles or fully automating some of them.

2.2.2 Intelligent digital twin and AI integration

Intelligent digital twin (IDT) refers to the digital twin with modelling and simulation enhanced with artificial intelligence (AI). This concept has also been labeled as the cognitive digital twin (Zheng et al., 2022) or the digital triplet (Alimam et al., 2023). However, the contemporary digital twin solutions consistently include some level of AI. AI utilizes Machine Learning and Deep Learning (ML&DL) to detect correlations and complex patterns from large amounts of data (Sahoo et al., 2023). This enables for example real-time anomaly detection, intelligent scheduling, and predictive maintenance. Grieves (2022) mentions four attributes that intelligent digital twins have as opposed to “traditional” twins: they are active, online, goal-seeking and anticipatory. Zheng et al. (2022) conclude that there have been multiple studies from recent years discussing the different possibilities, functions and definitions for cognitive digital twin. They discuss how the availability of real-time data allows for the automatic learning, reasoning and adjusting.

Front running simulations allow the IDT to notice issues in the manufacturing process before they happen (Grieves, 2022). Depending on the available data and the granularity that is applied, it is possible to prevent large scale complications or even future failures on specific tools, for example smart tools. AI models require training, and training requires data – for example past scenarios and their consequences and image recognition. Digital twin supports AI by providing accurate training material, like simulations and new product visualizations. Symbiotic Autonomous Systems (SAS) can use the information, simulations and predictions generated by digital twins to make smarter decisions, adapt to changing conditions and enhance overall performance (Chen et al., 2023). This synergy enables DT to use the real time data gathered by SAS sensors, robots or systems, and SAS to proactively address challenges in the physical world without human actions.

2.3 Digital twin in the manufacturing industry

According to Fortune Business Insights (2025), manufacturing sector is predicted to be the most expanding sector in the global market for digital twins. It is expected to increase from 16.42 billion euros (2024) to 240.11 billion by 2032 – a 39.8% growth rate annually (Fortune Business Insights, 2025). Additionally, 29% of manufacturing companies have executed digital twin strategies either fully or partially, while 65% of decision makers in this sector are planning to (Hasan, 2023). The fast-growing popularity of digital twin technology in the manufacturing industry suggests that companies that fail to adopt it risk falling behind in productivity, efficiency, innovation, and overall competitiveness.

2.3.1 Expected benefits

Anaya et al. (2024) discuss the benefits of a digital twin in every phase of manufactured object's life cycle. They divide the benefits into design phase, manufacturing phase, service phase, and disposal phase. Benefits of the design phase can be seen in resource planning as well as designing the product, while the focus of service and disposal phase is monitoring the fielded product and finally re-use and recycling opportunities support (Anaya et al., 2024). Manufacturing phase, which is the emphasis of this thesis research, can expect benefits in equipment maintenance, process optimization, quality management, supply chain management, operator training, and production planning (Anaya et al., 2024). Additionally, Gupta & Starr (2014) summarized that manufacturing operations consist of design, material procurement, production planning, manufacturing, quality control, assembly, testing and inspection, and delivery, and each of these phases can be optimized with a digital twin. Energy efficiency and decision-making impact are digital twin benefits that are not specific for the manufacturing industry, however they are valuable (Xia et al., 2025).

Both product and process design are less time consuming, provide more quality, and allow for better design-for-purpose solutions when they are completed in a digital environment (Attaran & Celik, 2023). Digital twins provide the ability to create virtual prototypes of products or manufacturing equipment to test different features and designs without the need to invest in multiple physical prototypes. Virtual process design can aid in new process development by allowing process optimization under different conditions and identifying inefficiencies before moving to the factory floor, minimizing disruption to production (Attaran & Celik, 2023). Also Abanda et al. (2025) emphasize the role of digital twins in the design phase and mention customized products and rapid, optimized design, as well as virtual testing.

Digital twin can aid material procurement in route efficiency, packaging performance, and just-in-time or sequential production optimization (Attaran & Celik, 2023). Distribution route analysis defines the optimal route according to parameters like cost and time (Attaran & Celik, 2023). Supply chain connectivity and transparency can be improved with a digital twin which delivers real-time data about inventory status, supplier performance, and lead times (Anaya et al., 2024). This enhanced visibility supports informed decision-making and fosters collaboration between different supply chain partners. Additionally, risk controlling, warehouse and storage planning, and route and location optimization are possible use cases (Abanda et al., 2025). Sales can utilize real-time product details, which reduces cost and time (Abanda et al., 2025).

Production planning and scheduling can be challenging in a complex manufacturing environment. Digital twin can take into account constraints such as current demand, resource availability, lead times, and capacity, and develop optimized production schedules (Anaya et al., 2024). In addition, any unexpected events such as machine failure can change production plans automatically by using real-time updates from the physical twin, and resource allocation problems caused by imbalance of information can be minimized (Abanda et al., 2025).

Digital twins are essential in supporting quality assurance and inspection within manufacturing. They can identify differences from the required specifications and recommend adjustments to maintain quality standards. This is enabled with a digital model of the product (Anaya et al., 2024). Additionally, the potential of automated inspection tasks help minimize human error and boost overall efficiency (Anaya et al., 2024). Real-time detection of defects, quality prediction, and process evaluation and parameter optimization is possible by utilizing digital-twin driven quality control methods (Abanda et al., 2025). Testing and inspection of products can be enhanced with dynamic cross-checking or most common assembly defects summaries. Additionally, equipment performance issues can be anticipated with digital twins before they occur by monitoring the data and comparing it to historical or simulated data (Anaya et al., 2024). This information helps minimizing unplanned downtime and allows for proactive maintenance. Digital twins are able to model both manufacturing processes and individual equipment, and identify variance which could indicate part or system failure (Attaran & Celik, 2023). Equipment diagnostics can also be conducted reactively after a failure has happened, which leads to faster fault analysis and resolution (Abanda et al., 2025).

Process optimization is possible with simulations and prescriptive analysis. By simulating different operational scenarios, digital twins can help optimize manufacturing workflows by identifying inefficiencies and bottlenecks in the production environment. Practicing assembly in a virtual setting can minimize errors and reduce the operator learning time. Digital twins can simulate the physical tools, materials, and environment, and enable customized feedback for each operator by assessing individual performance. (Anaya et al., 2024)

Digital twin enabled energy management system considers production load real-time data, operating parameters and conditions to reflect the production process energy consumption (Abanda et al., 2025), resulting in energy efficiency improvement. When real-time monitoring or energy usage is combined with bilateral data transfer, the shop-floor energy usage can be controlled remotely. Xia et al. (2025) designed this type of real-time controlling digital twin for the purpose of optimizing production line energy usage. Their decision model simultaneously considers machine utilization and maintenance, production volumes, and energy costs, making real-time optimization decisions and adapting to unexpected situations. Additional use cases for real-time control include, for example, re-programming of collaborative robots via digital twins (Rivera-Pinto et al., 2024). This user interface interaction has multiple different possibilities from machine speed and temperature adjusting to starting and stopping whole production processes (Rivera-Pinto et al., 2024).

With digital twins, manufacturing plants can transform their way of working from being reactive to being predictive. Data management and information tracing in complex manufacturing environments is essential (Abanda et al., 2025). Ready, processed information creates advantages for decision making by itself, but when enhanced with advanced analytics and for example prescriptive analytics, data-driven decision making is effortlessly accessible.

Benefits and fidelity are interlinked, and therefore fidelity must be considered early in the design and concept development phases. This ensures that the assumed benefits are achievable, since a more accurate and detailed system leads to a higher cost. In addition, lower complexity models are less time-consuming to create, require less data input, utilize fewer resources, and generally the interpretation and use are easier. In Figure 8, the net benefits are highest at optimal fidelity; when fidelity is higher than that, the relation of costs and saving potential is not optimal. (Kober, Algan, et al., 2023)

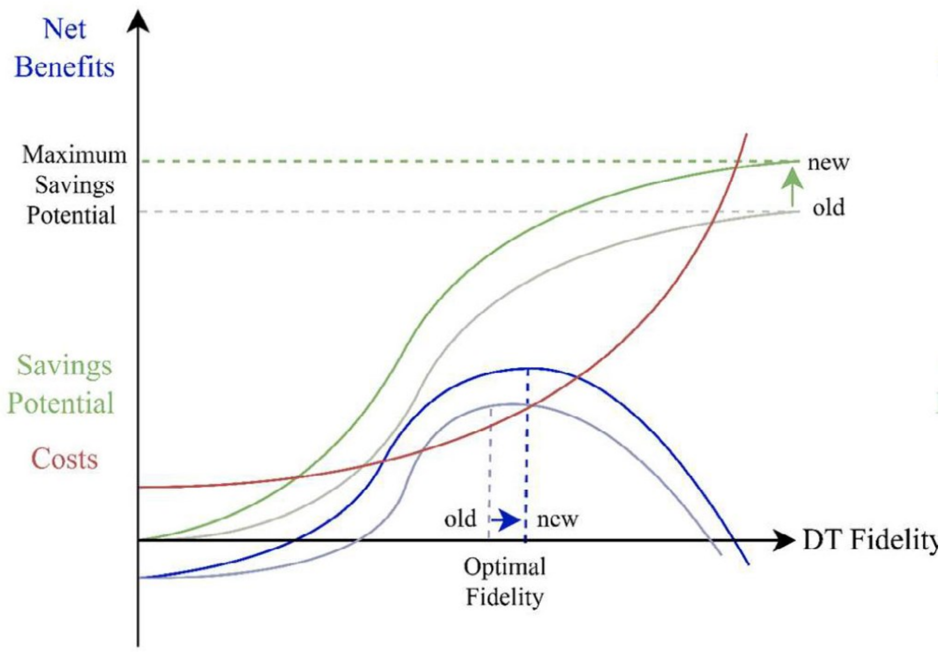


Figure 8. Rising savings potential (Kober, Algan, et al., 2023).

Rising savings potential has an impact on the optimal fidelity – when expected benefits are higher, the optimal fidelity is higher (Figure 8). Additionally, when costs of the system rise, the positive range of fidelity gets more limited. This graph creates understanding on the importance of well executed net benefits calculation of a digital twin solution.

2.3.2 Data collection and management

In the manufacturing environment, product quality is effected by manpower, machine, material, method, measurement, and environment, and all of these elements in the modern industry collect data (M. Liu et al., 2021). This enormous number of data points and sensors ensure that the problem with data is most often not the lack of it, but the inability keep up and extract valuable information from it. A digital twin that utilizes AI and DL has the ability to process raw data without manual feature selection and learn directly from the data, presuming that it has large datasets to learn from (Sahoo et al., 2023). In the manufacturing industry and its data overload, this can bring benefits in multiple different functions. For instance, the amount of data is usually very high in automation systems, which can create benefits in maintenance - the twin can process automation system datasets as they are created, find hidden patterns, and identify potential issues in the future (Sahoo et al., 2023).

Data collection is essential in the online, real-time requirements of a digital twin. If there are not enough data points currently in place, there are multiple ways to collect more data in the manufacturing environment. IoT generates data through sensors, and CPS monitor it by computing, communication and real time control (Li et al., 2024). Sensor can be defined as anything that converts the variables of the physical world into data. Some physical data collection methods include radiofrequency identification (RFID), scanners, cameras, sensors, and readers (Dihan et al., 2024). Additionally, data can be collected via software, for example with multi-modal data acquisition technology, value-stream mapping, augmented- and virtual reality, application programming interface (API), or software development kit (SDK) (Dihan et al., 2024).

Multiple data sources, forms, and scales in regard of volume and time propose a challenge for the development of a digital twin (Resman et al., 2021). For example, the data can be collected from MES, ERP, machine sensors, operator activity or production planning. Anaya et al. (2024) discuss the importance of adequate data collection platform; one important characteristics of such platform is the capability to collect and store different types of data from sensors, process parameters, environment and machines. Another characteristic is the administration of real-time data flow in large quantities and without delay. In order to reach these goals, the platform should operate via acknowledged industrial protocols, for example OPC-UA, MQTT or Modbus (Anaya et al., 2024).

Cloud services provide storing and managing of data in a centralized location, which makes it easy to access. Additionally, data processing and analysis tools, for example pattern and anomality detection are provided by cloud services. They make data sharing more efficient between different stakeholders and provide scalability for when digital twins evolve. Some examples of cloud services include Microsoft Azure, Amazon Web Service (AWS), Google Cloud, IBM Cloud, Oracle Cloud, and Alibaba Cloud. (Dihan et al., 2024)

2.3.3 Challenges and enablers

During the development stage, the challenges can be divided into development issues, system integration, security, and data quality issues. Complex manufacturing environment makes it challenging to create a model which is up-to-date, accurate, and reliable, while multiple different system properties and lack of standardization in languages and ontologies lengthen the development time and increase computational performance requirements. Old systems and equipment might need to be updated or integrated to new technologies, which both can be challenging and require investments. (Perno et al., 2022)

The ensuring of data quality consists of the acquisition of the right data, which is described to be a problem especially in production process modeling, and ensuring the data is reliable and accurate by data validation (Perno et al., 2022). In order to protect confidential company data from mischievous behavior and unauthorized personnel, robust security measures like encryption and data anonymization must be implemented (Anaya et al., 2024). Data transparency and security is possible to address for example with blockchain technology (Perno et al., 2022).

Performance issues, organizational decisions, and environmental factors are barriers that can occur outside of the development stage. Hardware and software limitations pose a challenge to real-time monitoring, because the combination of big data and real-time communication is difficult, resulting in performance issues (Zhang et al., 2017). Digital twin implementation and operation requires skilled workforce, which is costly and scarce in the state-of-the-art technology (Perno et al., 2022). The lack of standardized tools and methodologies, as well as education and literature immaturity, are all environmental factors which lead to the inability to make organizational decisions related to digital twins (Perno et al., 2022).

Computational cost and architecture requirements depend on system complexity, data frequency, IoT sensor count, cloud vs. on-premise solution, and the type of analysis (predictive/prescriptive/real-time control) performed. Capital expenditure (CapEx) includes compute resource, sensors, software licenses, metrology equipment, and engineer time. Operational expenditure (OpEx) costs include energy costs for running simulations, cloud storage and computing fees, and software updates. Cloud-based solutions are more cost-effective. (McClenaghan et al., 2023)

The computational requirements and cost are difficult to estimate, since the scope and complexity differ. For example, ML consists of thousands of data points and runs on CPU, while DL utilizes millions of data points and requires GPU (Thompson et al., 2020). GPU characteristic include FLOPs (Floating Point Operations, per second), power consumption, and memory size, which all affect cost (Desislavov et al., 2023). Similarly, different simulation models are highly different in terms of cost and complexity. In an experiment by McClenaghan et al. (2023), a Finite Element (FE) Model and an analytical model were compared for the same purpose of modeling a thermal test bed. In the FE model, the CapEx was calculated to approximately 700 pounds (including a fusion subscription of 500£ and expert time roughly 35 £/h), while OpEx as 189 pounds to monitor one simulation run in the cloud with Fusion 360. This was calculated with the average simulation time being 352 seconds and 21 simulations required to cover an entire run. In comparison, they conducted the same experiment with an analytical twin, which had the response time of 0.05 seconds and resulted in CapEx and OpEx total as less than 15£. However, the analytical twin had lower prediction accuracy – the mean deviation was increased by 28% and range of deviations was increased by 73%. Table 1 presents the comparison of these two models, comparing the setup time, prediction accuracy, CapEx, OpEx, and time lag.

Table 1. The comparison of the FE and the analytical twin cost and characteristics (McClenaghan et al., 2023).

Entity	Setup [h]	Accuracy [Δ mm]	CapEx [£]	OpEx [£/run]	Time lag [s]
FE	3	0.046	700	189	352
Analytical	< 1	0.059	< 10	< 5	0.05

This experiment was conducted to a thermal testbed which requires very high accuracy and is not strictly applicable to larger scale production environments. In addition, analytical models are unlikely to be a viable option for more complex subjects, such as shopfloors, because the required models might be exceedingly complicated (McClenaghan et al., 2023). However, the study demonstrates the differences of accuracy, time lag, and cost in different digital twin technologies, and emphasizes the importance of system comparison and verification as well as the cost and benefit evaluation prior to investment.

Even though the novelty of this technology makes the development challenging, Industry 4.0 provides multiple enablers. AI makes it possible to realize the full capacity of digital twins by analyzing large amounts data without the need for identifying and applying complex rules to the analysis process. VR enables remote assistance, training, and simulations in a virtual environment, resulting in a more interactive usage of the technology. IoT/IIoT facilitates device communication and makes it possible to collect production data in large volumes, and development technologies, such as simulations, enable higher model accuracy and validation. (Perno et al., 2022)

Additionally, data management is made accessible with 21st century technologies. Edge computing, cloud computing, centralized data bases, and communication protocols such as MQTT, OPC-UA, and MTConnect ensure the availability, quality and security of data. System integration is possible to navigate with new IIoT solutions and platforms which can be designed to be compatible with current systems. In addition, protocols like OPC-UA are vendor-neutral and allow the communication between new IIoT and older systems. (Perno et al., 2022)

Knowledge building, in particular the Industry 4.0 standards and workforce training are able to reduce environmental issues. The usage of standards creates the foundation of data framework in enterprises, which then allows reliable data exchange between the virtual and physical twin. Adequate training of workforce will generate qualified decision-makers and ensure the expertise will not be upon few individuals and roles. Knowledge-building technologies, such as generating new knowledge through troubleshooting analysis and production-, material-, and sales data analysis, can be used for example in production optimization and new product development. Design considerations can similarly remove organizational barriers by ensuring that the models are scalable, interoperable, and in optimal fidelity, which creates a digital twin that can be used as a baseline for the creation of similar asset models. Hardware availability and the decrease of hardware costs can reduce performance issues. The latest hardware solutions increase computational power, which enables advanced machine-learning models and simulations. (Perno et al., 2022)

2.3.4 Comparison of existing solutions

D. Lehner et al. (2022) evaluated three digital twin platforms, Microsoft Azure, Amazon Web Services, and Eclipse ecosystem, to determine the differences between the platforms regarding 13 different requirements. The study concluded that reusability, security, and interoperability were supported by all the platforms, while automation protocols and convergence were not sufficiently supported by any of the platforms. There were slight differences between platforms in validation and verification, system interoperability, CI/CD, and domain expert involvement (D. Lehner et al., 2022). The platforms are constantly evolving, and even though the research does not describe the current state of these platforms, it highlights the difference of features between solutions. The most popular solutions were compared by Daase et al. (2023). Table 2 presents Siemens, Microsoft, Oracle, Ansys, and SAP solutions and their features.

Table 2. Comparison of popular digital twin solutions (Daase et al., 2023).

Provider	Solution	Features
Siemens	Plansight	Focused on manufacturing process. Combining data, contextualizing it, and providing visualizations.
	Simcenter	Focused on design and developing of products. Combining data, contextualizing it, and providing visualizations.
	Teamcenter	For product lifecycle management, except for usage and disposal data. Integrates product information from different departments and domains.
Microsoft	Azure Digital Twin	IoT focused solution. Representation of processes or things of the physical world.
Oracle	IoT Digital Twin Simulator	Simulation models of manufacturing environment that are ready to use; includes alerts.
Ansys	Twin Builder	Product focused, includes predictive maintenance capabilities.
SAP	SAP Predictive Engieering Insights	Monitors and simulates past, present, and future. Aim is to increase industrial assets' overall effectiveness.
	SAP S4/HANA	Includes modules, such as SAP Advanced Planning and Optimization (APO) which has the capability to forecast, do production planning, and manage orders and distribution, and SAP Enterprise Asset Management (EAM), which is created for physical entity management.
	SAP Landscape Transformation Replication Server (SLT)	Facilitates real time data exchange to third party cloud service environments like digi-tal twin platforms.

2.4 Design and development

The most crucial stage of implementing a digital twin in an organization is defining its purpose. This stage provides answers for questions like what problem is it solving and what will the digital twin represent. After that is defined, the implementation consists of six consequent steps, which are 1 - create (sensors, process and environment inputs), 2 - communicate (network communication), 3 - aggregate, where data is sent to a data repository and prepared for analyzing, 4 - analyze with advanced analytics to identify improvement opportunities, 5 - insight (decision making models created), and 6 - act, where recommendations are fed back to physical system (Kuehn, 2018). The development of a digital twin should be approached from a systems engineering perspective.

Without the use of model-based systems engineering (MBSE), the software may experience poor scalability, high maintenance and upgrade costs, and low reuse rate of system blocks (J. Liu et al., 2021).

2.4.1 Requirements management

Requirements management or requirements engineering is the practice of identifying, analyzing and implementing user needs (Nuseibeh & Easterbrook, 2000). The Chaos Report by Standish Group (2020) concluded that only 31% of software projects are completed fully according to the specifications. According to Nuseibeh & Easterbrook (2000), the five core activities of requirements management are

1) Eliciting
2) Modelling
3) Communicating
4) Agreeing
5) Evolving,

not used in that sequence, but rather iteratively. Also Pandey et al. (2010) propose a similar four-phase process consisting of elicitation and development, documentation, validation and verification, and requirement management. Eliciting refers to the activity of collecting information from stakeholders in order to create requirements and thus achieve system boundaries (Nuseibeh & Easterbrook, 2000). In addition to users, stakeholders refer to the customers and developers of a system. Elicitation relies mostly on communication, which makes it vulnerable to language barriers and misunderstandings for example between software engineers and management. Users can adopt different scenarios of working with and around the system when communicating their needs; these are called use cases (Nuseibeh & Easterbrook, 2000). They allow for more specific communication of each need that the system needs to fulfil.

Conflicting interests between stakeholders can make agreeing of requirements difficult. Nuseibeh & Easterbrook (2000) propose multiple solutions to managing this validation phase. They state that evolving requirements usually consist of fixing errors, deleting obsolete requirements and adding new requirements. Change management, traceability, inconsistencies, and development of product families create the need for flexibility in the design process. The V-model of systems engineering (Figure 9) is used for managing complex system development, because it emphasizes testing and verification at every stage of development (J. Liu et al., 2021).

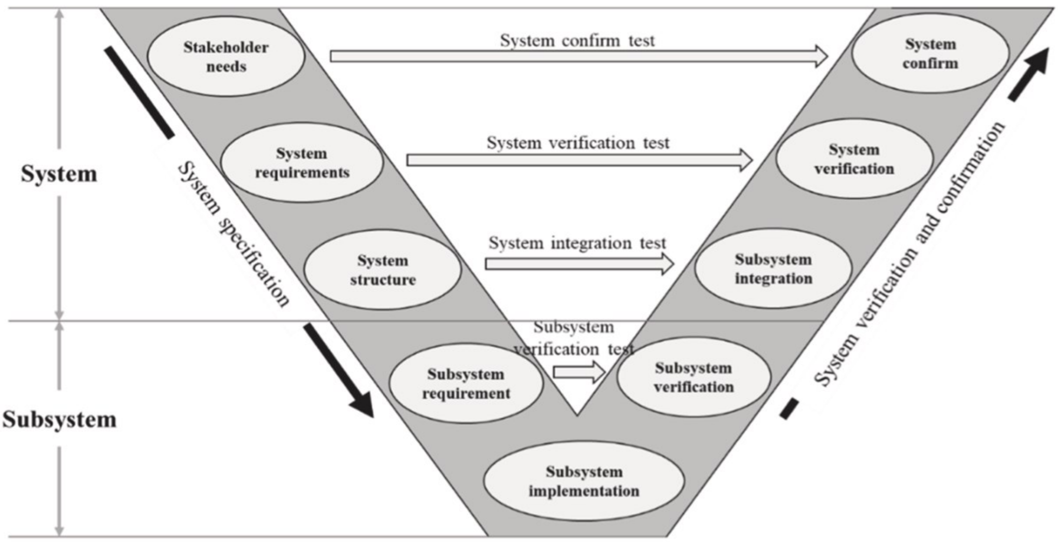


Figure 9. The V-model of systems engineering (J. Liu et al., 2021).

System requirements can be divided into three main categories, which are business requirements, stakeholder requirements, and solution requirements. An example of a business requirement is "we have recognized a need to see the production process in real time to optimize and manage...", describing the reason behind system implementation. Stakeholder requirements include the views that future users have for the system based on their specific role. For example, "as a project manager, I need to forecast..." or "as a production development engineer, I want to control…" are needs that define the user and validate their requirement. Solution requirements are divided into three types: functional requirements, non-functional (quality) requirements and technical requirements. Functional includes a product characteristic for every stakeholder need – for example, "the user can see the downtime of each piece of equipment". Non-functional or quality requirements describe the acceptable quality standard of each need. An example of a quality requirement could be “the system must run smoothly with 500 simultaneous users”. Technical requirements explain how the functional and non-functional needs can be met. This includes identifying the necessary data sources and data management. An example of a technical requirement would be “every piece of equipment must have a data collection log that is connected to the server”. (Pandey et al., 2010) (J. Liu et al., 2021)

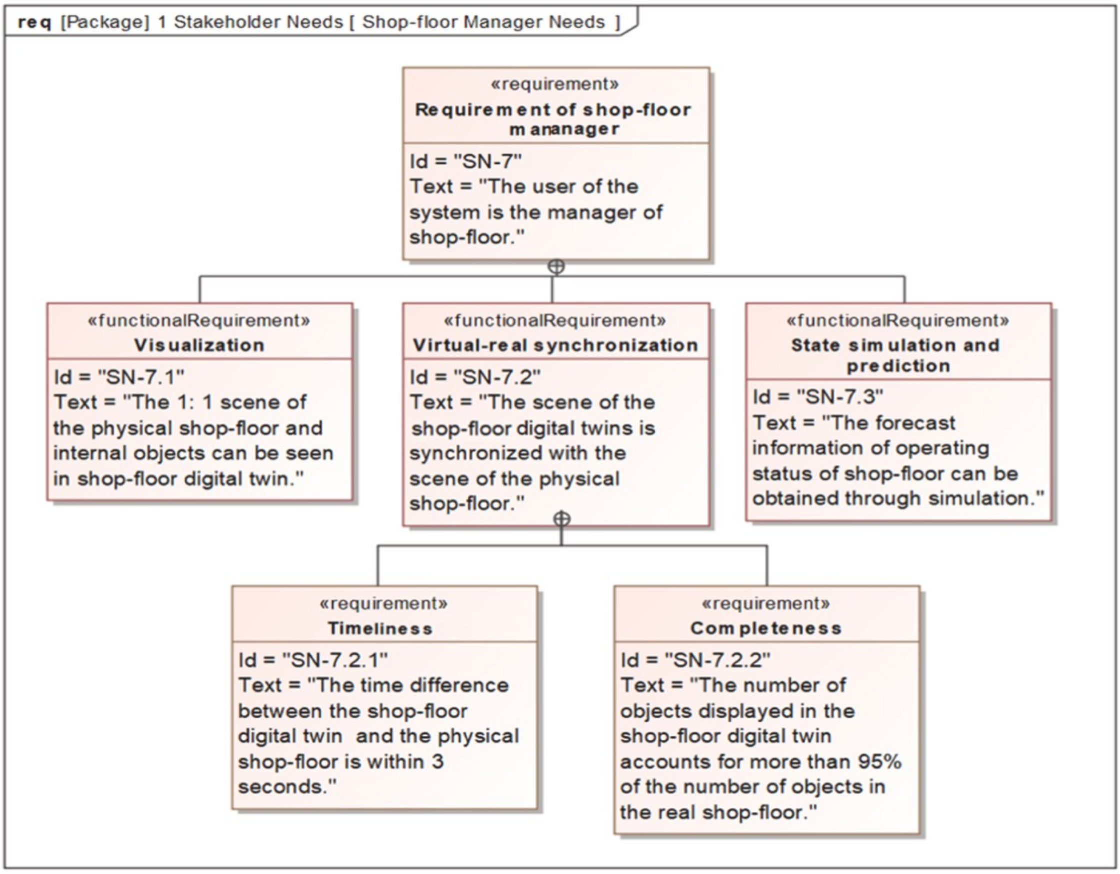


Figure 10. Example of a solution requirement diagram (J. Liu et al., 2021).

Figure 10 represents an example of a solution requirement diagram, describing the shop-floor manager’s functional requirements as visualization, virtual-real synchronization, and state simulation and prediction. Additionally, virtual-real synchronization creates two quality requirements in this example, timeliness and completeness, which describe the system's quality standard by defining the acceptable time lag to 3 seconds and the precision to >95% of real shop-floor objects.

2.4.2 Development process and architecture

To create a cohesive model for this development phase, Resman et al. (2021) introduce a five-step approach (Figure 11). In addition to describing the creation process of a digital twin, this model also represents the different types of digital twins. The outcome of step 2 is a digital model, step 3 represents a digital shadow with only parameters from the physical world (one-way), and step 4 is the final two-way communicating and online system, digital twin. Step 5 focuses on user experience by allowing every user to modify their experience and filter excessive information while focusing on relevant data.

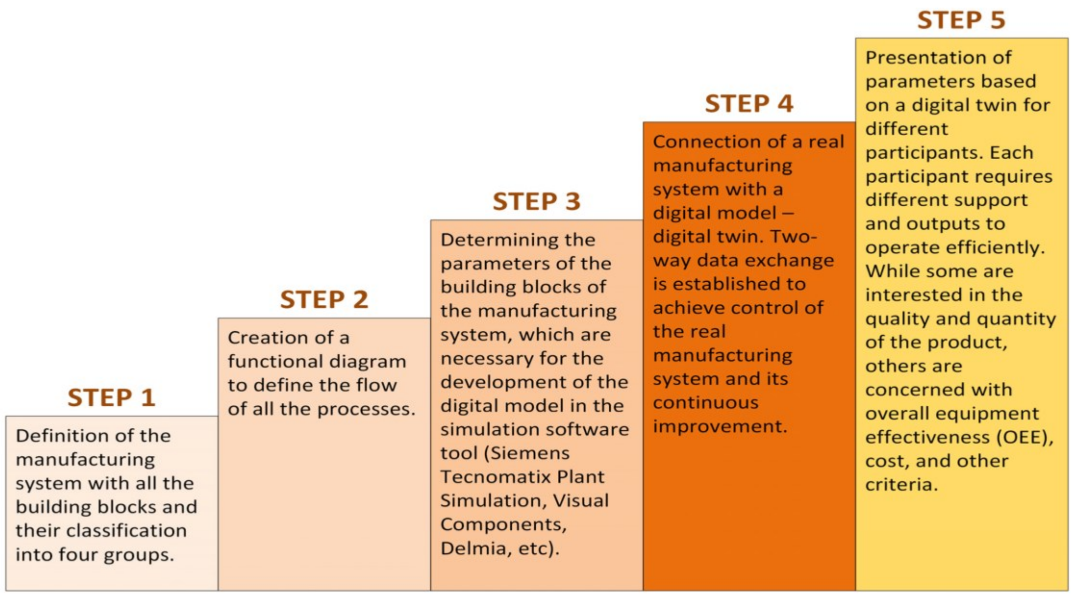


Figure 11. Five-step approach for establishing a digital twin (Resman et al., 2021).

The basic architecture of a digital twin is presented in Figure 12. It consists of physical world sensors which collect data through IoT communication and store it to the cloud. Data is then analyzed with AI and ML, which provide analytics and for example predictions. This is then translated back to the physical world as a diagnosis, and results in diagnosis or updates. (Ibrahim et al., 2023)

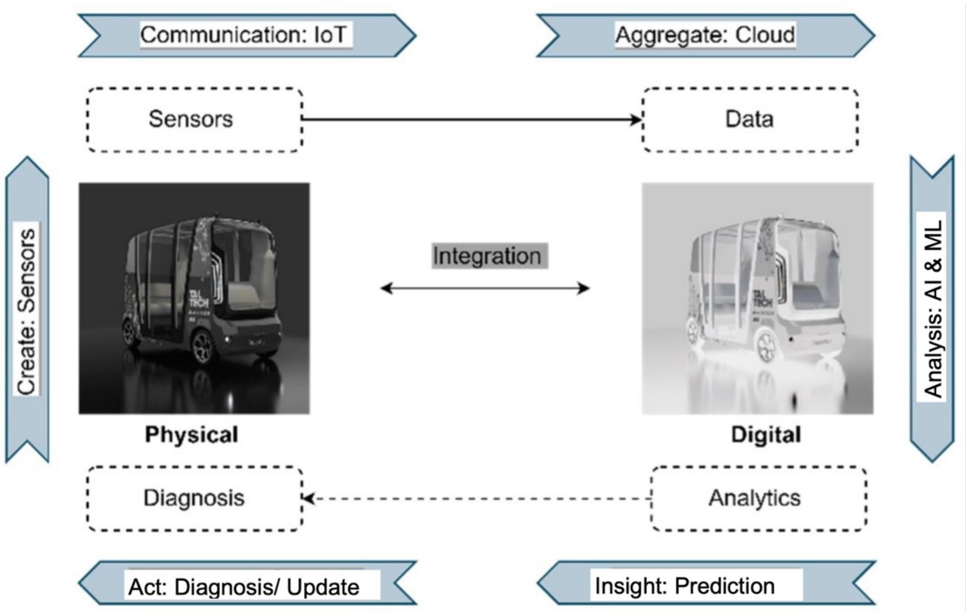


Figure 12. Basic digital twin architecture (Ibrahim et al., 2023).

The architecture is commonly divided in two main categories, which are model-based and data driven digital twins. Model-based is based on mathematical equations, while data-driven utilizes sensory data (black box). Thus, model-based is more expensive to create and maintain, but the time of creation is shorter. It is commonly utilized in model-lable physical systems, such as specific tools or machines. Data-driven is less expensive but takes longer to create, and is utilized in complex and cyber-physical systems, that are difficult to model due to noisy data and unknown variables. It requires substantial amount of data and is suitable for complex manufacturing systems and shop floors. (Ibrahim et al., 2023)

Digital twins are generally built using six components: Process Analytical Technology tools (PAT), control models, process models, information technology and operational technology architecture (IT-OT), user interface, and validation (Phalak et al., 2023). PAT framework allows the real-time data collection and measuring process evolution. Figure 13 represents the data flow between the different components. Data is first created from equipment with sensors (a), and it travels through offline data pipeline to data analytics platform (e), and through online data pipeline to the digital twin (c). Data analytics platform includes historical, simulated, and process metadata, and it constantly creates new models. It feeds the digital twin with validated new models, which digital twin connects with real-time sensor data from the online data pipeline. This results in predicted future states and optimized processes. Finally, the twin utilizes that information and is able to optimize the equipment via direct control and a control system (d1, d2).

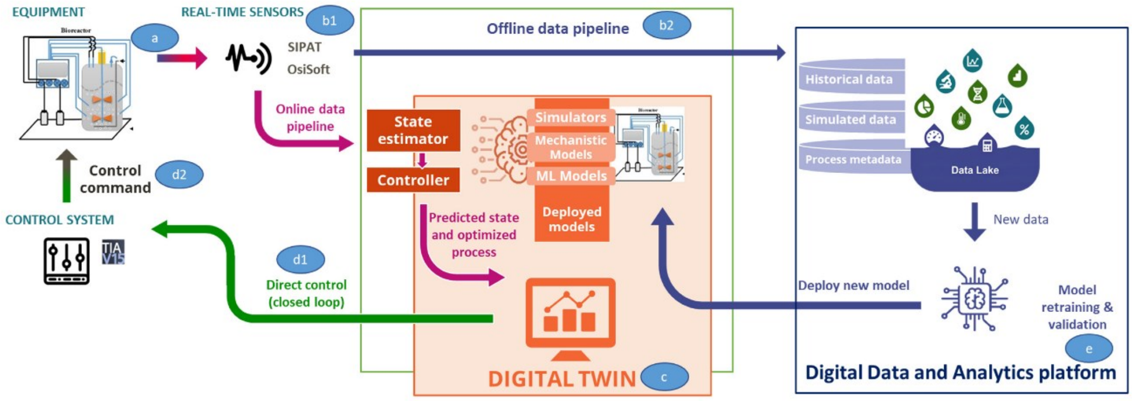


Figure 13. Data flow between different digital twin components (Phalak et al., 2023).

IT/OT converge enables the right information at the right time in the right format to every device, sensor, and person. OT, which includes advanced physical technologies (robotics, AGVs, etc.) provides data for IT which analyzes the data with advanced tools, such as learning algorithms and scenario analysis. The underlying technologies like platforms, security, and software are bringing OT closer to IT, while IT supports OT by not only supplying information, but also in enterprise architecture, software configuration actions, and building standards. The prerequisites for IT/OT converge are IT governance (managerial control and responsibilities pattern within organization), interoperability (linking systems), collaboration between IT and OT staff, and IT infrastructure flexibility, which includes platform technology, software applications, data, telecommunication, and network technologies. (Ehie & Chilton, 2020)

2.4.3 Shop-floor digital twin development case

The shop-floor digital twin can be described with four different components: Physical shop-floor, virtual shop-floor, shop-floor service system, which is also called the application layer, and shop-floor digital twin data (Corallo et al., 2021; Lattanzi et al., 2021). Physical shop-floor includes all factory resources, e.g. machines, materials, and operators, while the virtual contains the descriptor models (geometric, attribute, behavior rules, data fusion) of these resources (Lattanzi et al., 2021). For the development, Corallo et al. (2021) proposed a hexadimensional shop-floor digital twin or HexaSFDT. It includes a digital space consisting of model, data integration, and knowledge layers, a physical space represented by the physical layer (objects, process, and people), and application layer referring to the tools and techniques supporting the shop floor digital twin functionalities (Figure 14). The bidirectional data flow is enabled by a network layer consisting of CPS and communication technologies.

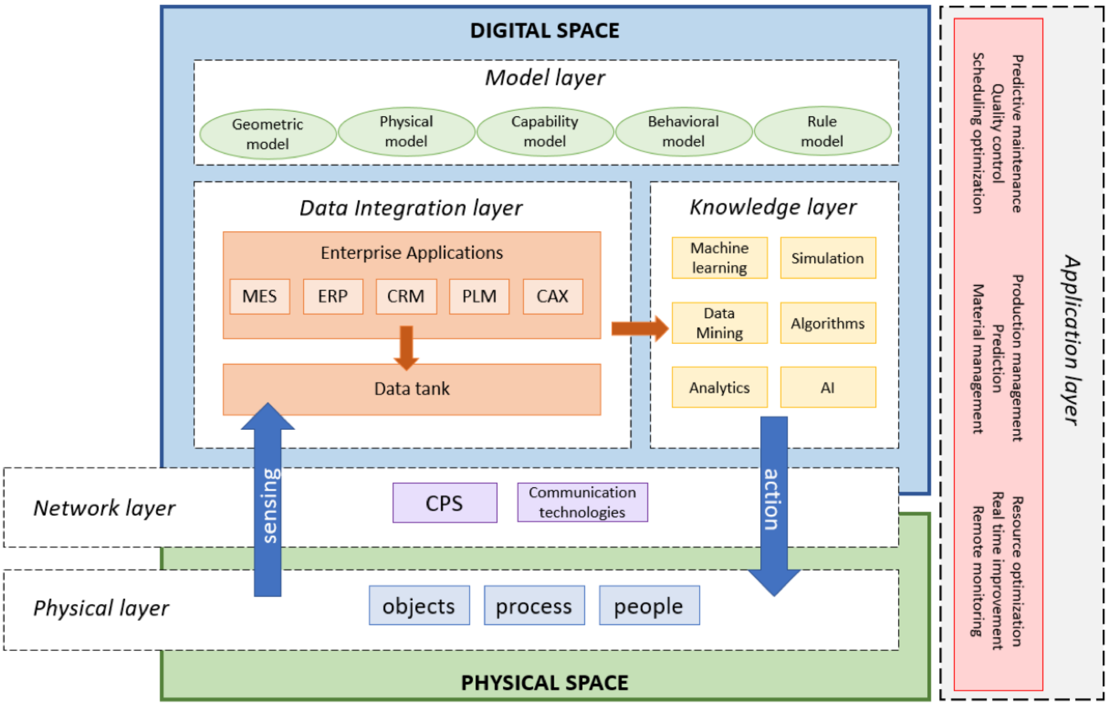


Figure 14. A comprehensive conceptual shop floor digital twin framework, HexaSFDT (Corallo et al., 2021).

J. Liu et al. (2021) created a shop-floor digital twin using modular, MBSE-based development. They chose MagicGrid modeling method because it is compatible with SysML. It divides the construction process into problem domain (stakeholder needs), solution domain (system architecture), and implementation domain (realization process). Visualization system is created using AutoCAD for layout, Creo and SolidWorks for product and equipment modeling, UG and SketchUp for model analysis, and 3DS MAX for the format conversion and lightweight. Rhino was used to model non-product and -equipment objects and Mixamo Fuse for human modeling. Finally, these visual models were imported to Unity3D to achieve acceptable level of visualization.

Synchronization system includes the realization of human/equipment/product/kanban state machine. The human state is represented by states like “idle”, “walking”, and “operation”. The real-time status is collected and stored in MES, and the real-time location data of personnel and material are collected with Ultra-Wide Band (UWB) and RFID tags and stored in a shop-floor IoT system, in UWB database. The status and location are integrated with data integration interfaces. This IoT-based data communication scheme is shown in Figure 15.

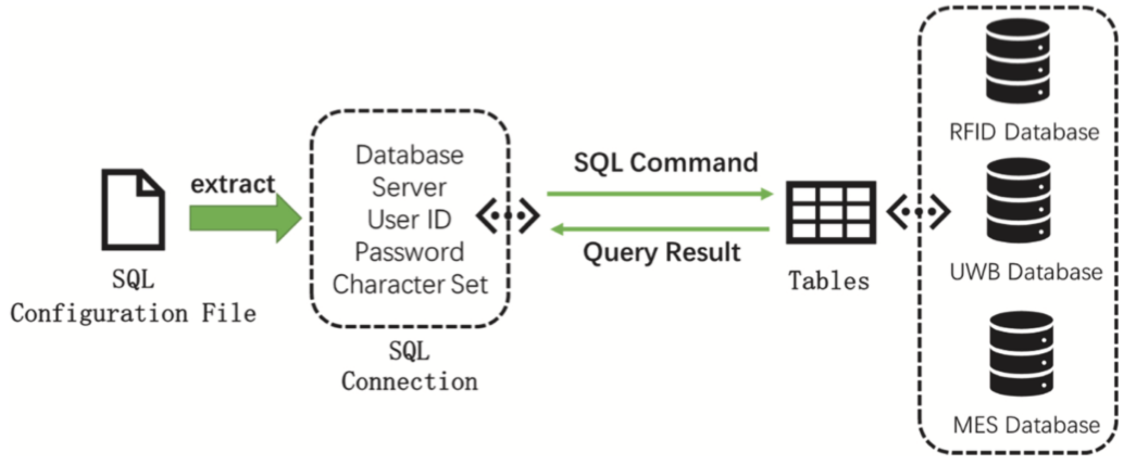


Figure 15. IoT-based data communication scheme for synchronization system (J. Liu et al., 2021).

After completing the data communication scheme, a simulation system was realized using the event-driven scheduling method which is based on the Future Event List (FEL). First the operation logic and production processes are analyzed and described with models, and events' distribution parameters are estimated using historical data. Then the simulation samples generated from the previous step are screened against actual floor conditions and simulation requirements, to dictate a sample input generator. Finally, the different simulation event processing logic is constructed, and it is triggered with the progress of the simulation clock. This creates a simulation which includes all current and future event description and their time of occurrence. The transient and continuous real-time simulation is represented in Figure 16.

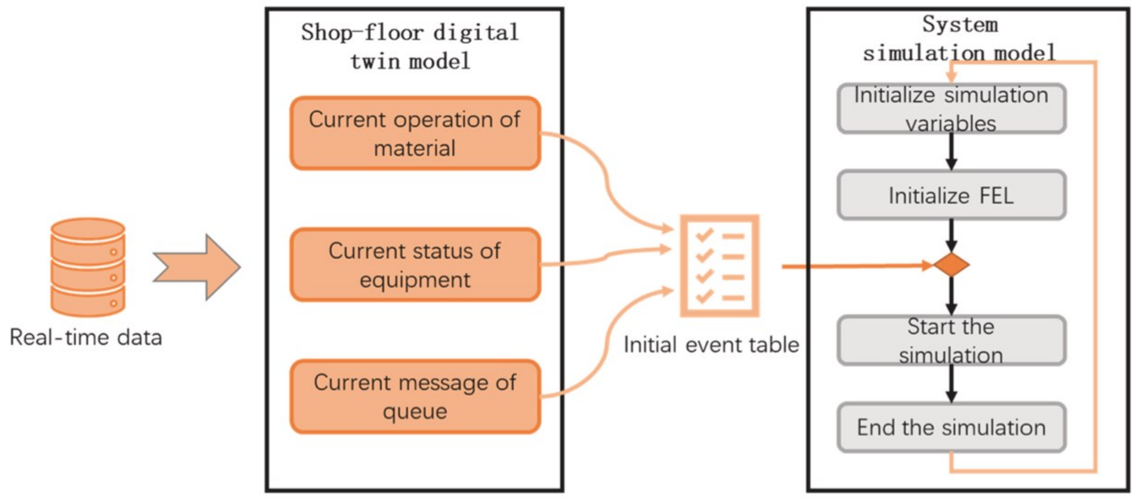


Figure 16. Continuous transient simulation based on real-time data (J. Liu et al., 2021).

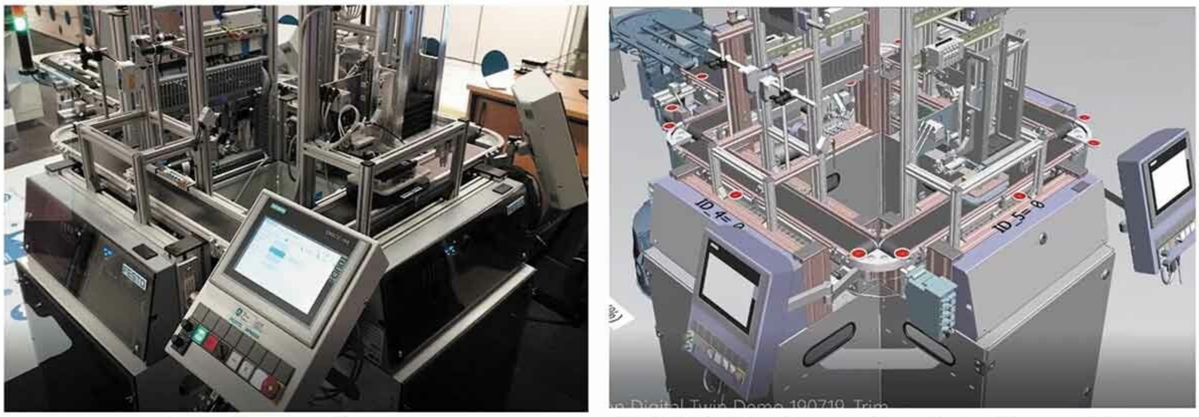
The advantages of modular development are customized requirements, which lead to the avoidance of top-level design errors, and a standardized library of models, which reduces software development time. Forecasting can be done either with performing big data analysis on shop floor performance, or analysis on offline simulations of dynamic variables, such as workpiece arrival and completion time. (J. Liu et al., 2021)

2.5 Industrial applications and case studies

Digital twins can be utilized in different purposes depending if the focus is on design, manufacturing, or operational phases. In manufacturing phase, the digital twin can be applied in real-time monitoring, production control, workpiece performance prediction, human-robot collaboration & interaction, process evaluation & optimization, asset management, and production planning (M. Liu et al., 2021). This chapter introduces five case studies of digital twins, their implementation process, and benefits. The use cases entail smart factory digital twin for process monitoring and skill development, shop-floor monitoring and prediction twin, automotive production line, two smart factory cases, and lastly solution demonstrations by two vendors, Process Genius and Roima Intelligence.

Case 1: Festo cyber-physical smart factory

Onaji et al. (2022) present a case study of a cyber-physical smart factory. The digital twin solution was created for teaching and developing automation-based skills in smart industrial environment. In addition to providing a learning environment, this system supported experimental analysis and process monitoring. Physical twin feeds the virtual twin with real-time data via RFID technology and sensors, which allows bottleneck analysis, and analysis of the impact that bottlenecks have on the process outcome.

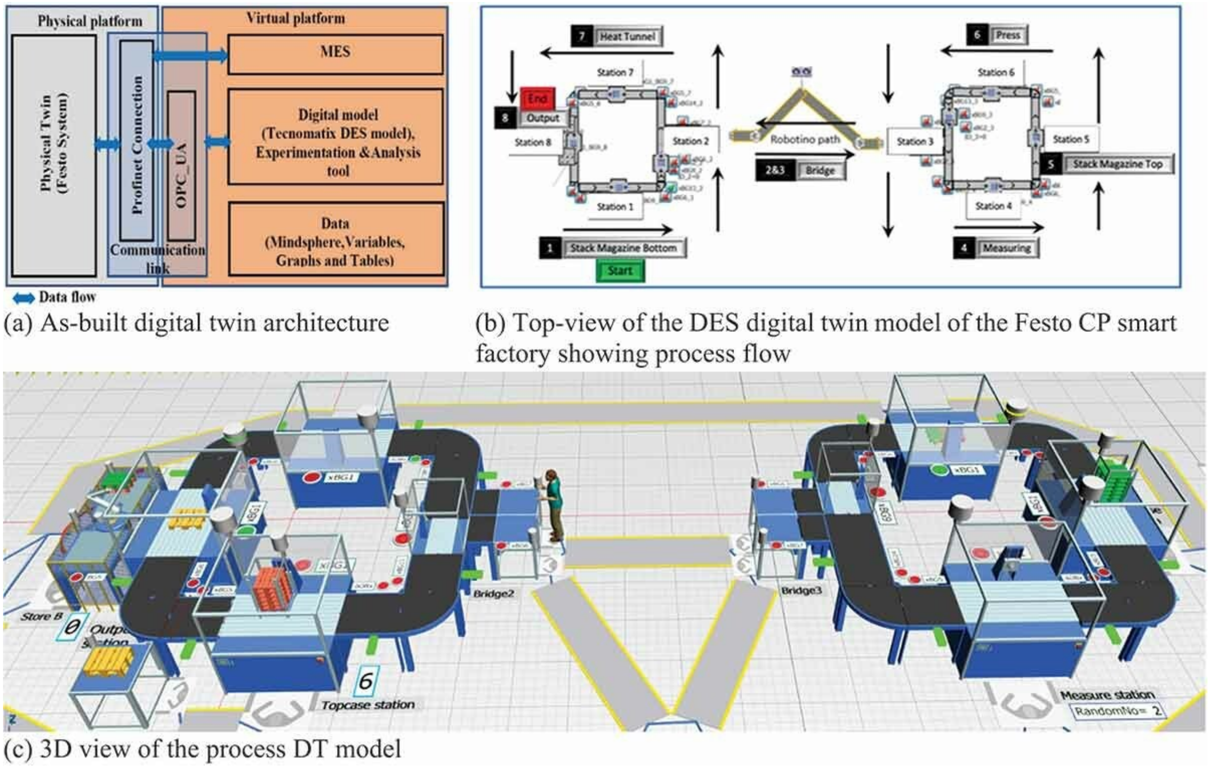


(a): Festo CP smart factory (b): DES digital twin of the Festo CP smart factory

Picture 1. Festo smart factory, physical system and its virtual counterpart (Onaji et al., 2022).

The system was managed through MES, where production orders are made. It consists of a production line with six work stations, transfer bridges, and one autonomous vehicle. Each module has their own PLC (Programmable logic controller). The model is three-dimensional and has high visual granularity, which supports its purpose as a learning tool. This level of detail can usually be achieved with laser scanning.

The development process was conducted in six phases: (1) conceptual model design, (2) development of virtual models, (3) control panel, (4) data layer, (5) intelligent layer, and finally (6) validation and verification process. The virtual models were created with Siemens NX and Tecnomatix plant simulation, which was chosen for its visualization and analytics capabilities and OPC-UA protocol. The twin is presented in Picture 2, where system architecture (a), process flow (b), and the final, 3D process digital twin (c) are displayed.



Picture 2. The modeled process flow in Festo smart factory (Onaji et al., 2022).

The completed digital twin identified bottlenecks for example in carrier-conveyor friction, which increased travel time and delay in finished products' unloading station. This information enabled the determination of optimal configurations and diagnostic analysis. In addition to process analytics, the twin had the ability to simulate modification decisions beforehand.

Case 2: Shop-floor monitoring and prediction system

Zhuang et al. (2021) conducted a case study of the construction and application of a digital twin a manufacturing company's shop floor. The company produces structural plates, and a digital twin-based visual monitoring and prediction system (DT-VMPS) of shop floor was deployed to enhance production efficiency, specifically to tackle on-site requirements response rate and production deviances. IoT for real-time data collection of resource and product locations as well as status of AGVs, robotics, and other equipment was already deployed.

The solution consists of a visual module for monitoring and showing prediction results, and a prediction module. Additionally, the solution includes integration with MES and a virtual scene in 3D. The chosen platform is Unity3D. It monitors equipment status and utilization, order status, product quality, and presents messages created by other platforms (Picture 3). This enhances transparency of production and improves response times to disruptions and other requirements. The case study concluded that the consistency of state between the physical and virtual was >95%, with a time delay of less than three seconds and model integrity rate as 97%.

Picture 3. Operation interface (Zhuang et al., 2021).

The construction of the system was done in four phases, building the digital twin layer by layer:

1. Modeling objects (production elements and processes)

2. Modeling dimensions (five-dimensional architecture)
3. Real-time status monitoring
    a. Data storage
    b. Data communication
    c. 3D visual monitoring of shop floor operating status
4. Status prediction
    a. Process performance prediction
    b. Data prediction of indicators for production elements
    c. Learning mechanisms of Markov chain model.

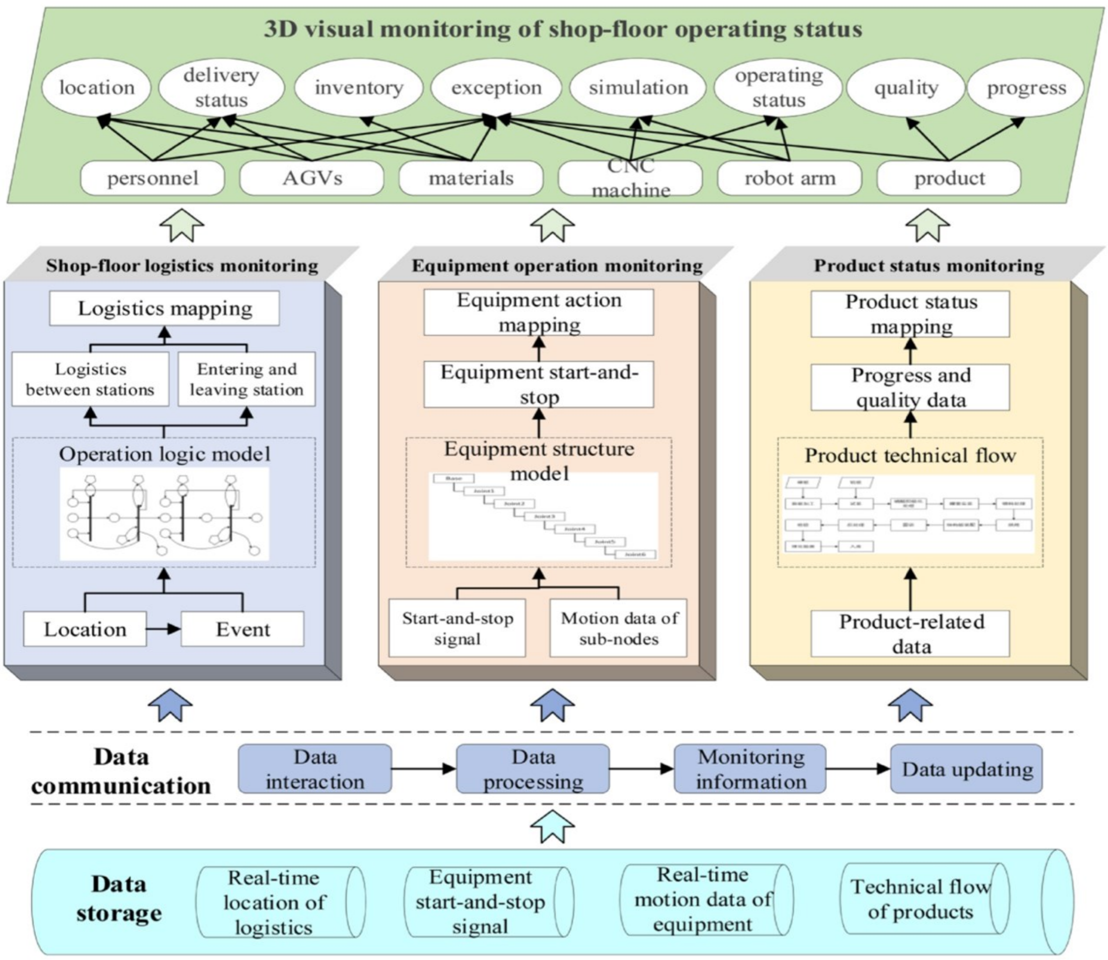


Figure 17. Implementation process of a digital twin -based 3D visual and real-time monitoring of shop-floor operating status (Zhuang et al., 2021).

The architecture is five-dimensional. To adequately describe the complex shop floor, there are five attributes that need to be considered: geometric, physical, behavior, rule, and data dimensions. After determining these dimensions, the status monitoring could be achieved. Status monitoring layer was created in three phases: data storage, data communication, and finally a 3D visual monitoring of shop floor operating status (Figure 17). The Markov chain method utilizes both parameter and structure learning. In the fourth phase, a predictive layer was developed using a Markov chain model. The learning models collect operating status dataset and update the Markov chain parameters accordingly. The complete integrated business workflow including MES, visual monitoring module, prediction module, and the 3D virtual scene is represented in Figure 18.

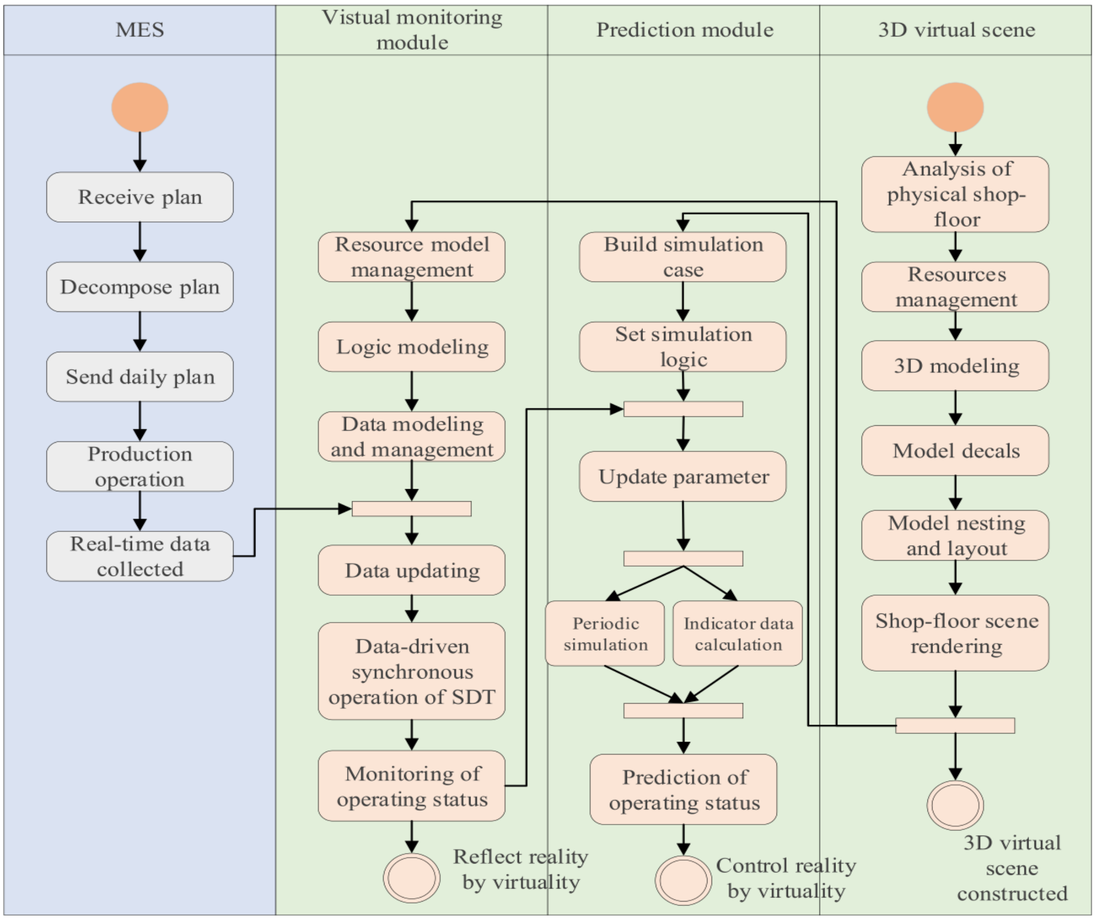
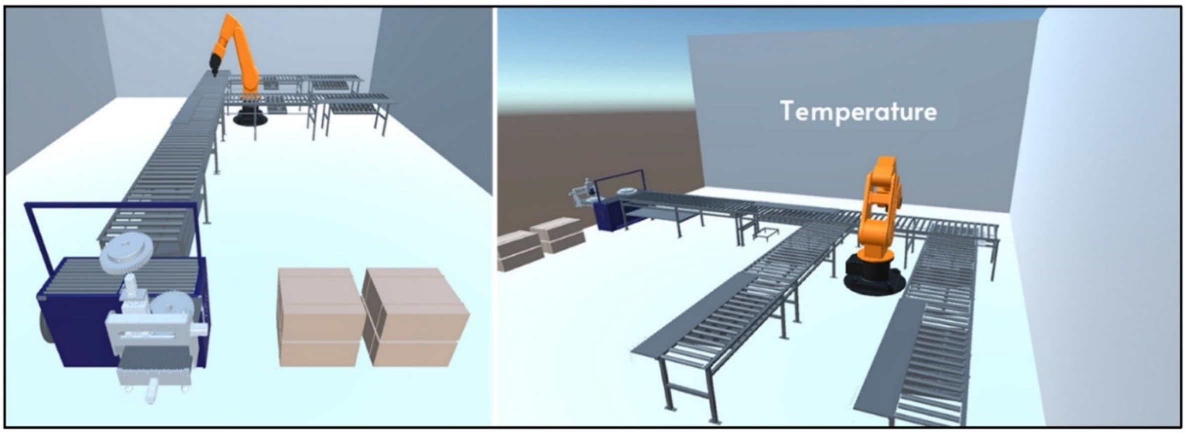


Figure 18. Integrated business workflow of DT-VMPS (Zhuang et al., 2021).

Case 3: Automotive production line

Mendi (2022) created a digital twin for automotive production line monitoring. The goal of the system implementation was to ensure that the production line operates at optimal condition, make analyses of the produced materials to enhance durability, enable instantaneous response to disturbances, and test variable scenarios. The finished product is a decision support system.



Picture 4. Process model of an automotive production line (Mendi, 2022).

Even though the created system was able to interfere with the physical system, the manufacturers preferred to have feedback that produces warning messages and leaves the action to humans instead of automatic intervention. Temperature-, dust density-, and frequency sensors were installed to gather data from the physical system. Optimal input ranges were defined, and the current state of the production line was presented in a visualization. This enables the operator to quickly notice when corrections must be made. The finished product is shown in Picture 4.

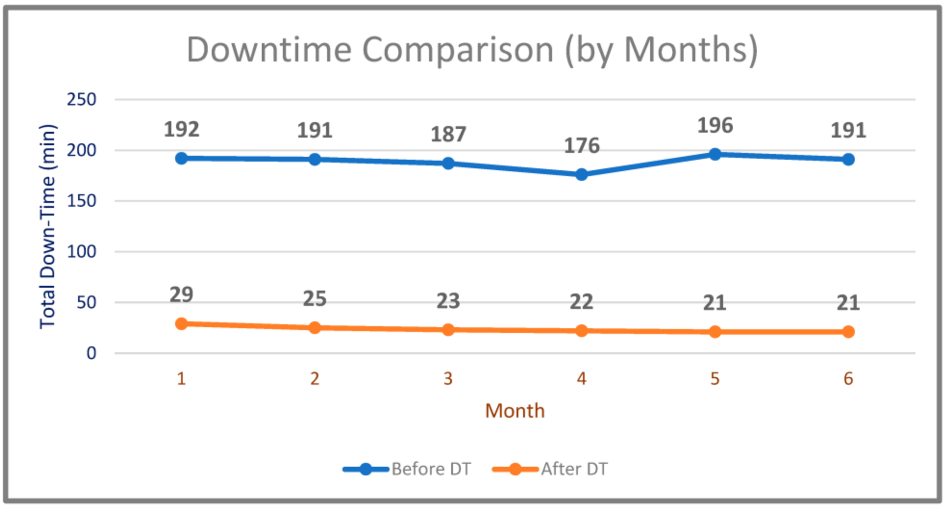


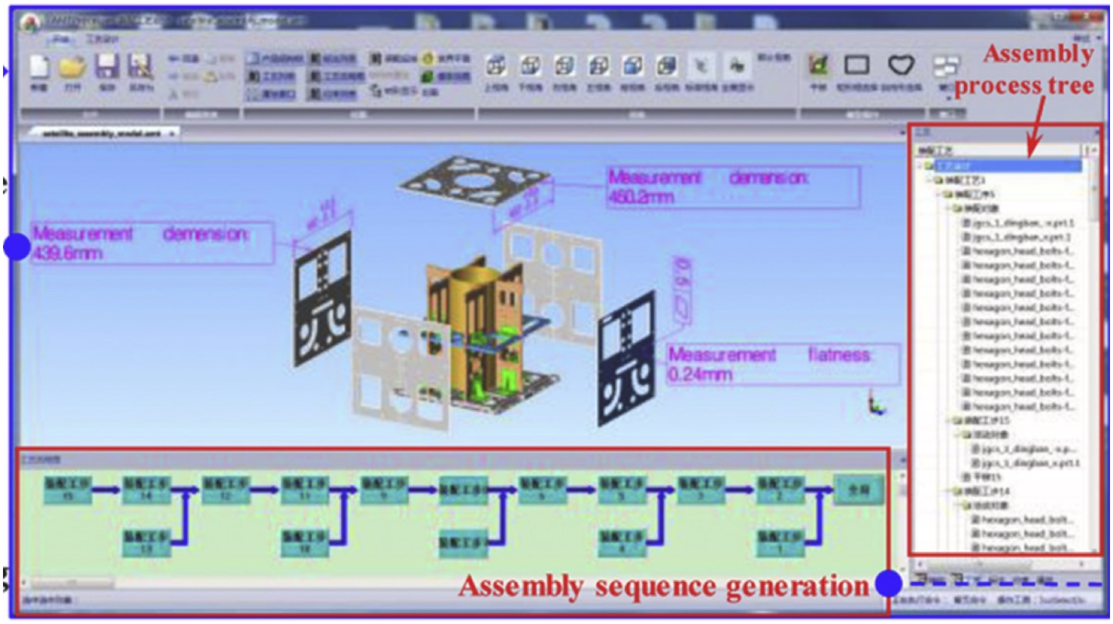
Figure 19. Downtime reduction after digital twin implementation (Mendi, 2022).

Following the digital twin implementation, the production line downtime was reduced by approximately 87% in a six-month follow-up period (Figure 19). Additionally, the daily machine working time (time required to do the same amount of work) decreased, leading to approximately 6% gain in efficiency compared to pre-digital twin.

Case 4: Satellite assembly: experimental test bed

Yi et al. (2021) created an experimental test bed for satellite assembly process. The process includes large number of variating, large scale parts, that are complex-shaped and require high precision. The assembly work is manual and is conducted in small batches or singles. These characteristics lead to assembly errors, complicated commissioning and difficulties in ensuring first-pass yield.

Digital-twin based assembly application system (DT-AAS) create an assembly model, sequence plan, assembly in-process model, simulation & verification of the process, and finally a tailored instructions output. This is enabled by feeding constant data from assembly process to the process design phase. By utilizing assembly precision predictions and experience methods generated during execution, constant improvement of work instructions is achieved.



Picture 5. 3D CAAPP process model for assembly process planning (Yi et al., 2021).

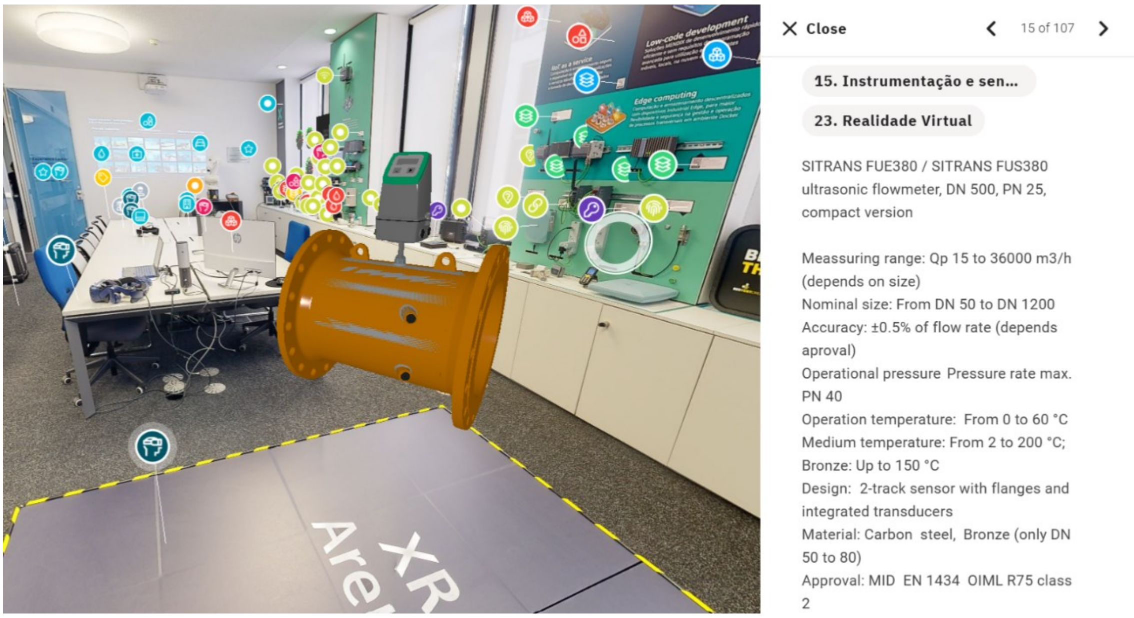
Following the implementation, First-pass yield (FPY) was improved by 21.4 %, reaching 96,4%. Additionally, the assembly commissioning time decreased by 53.8%, equipment accuracy improved by 62%, and rework count decreased by 83.3%. However, these exceptionally high numbers were achieved due to the complex nature of the original assembly process. Before the assembly simulation of specific products and tailored instructions, the process was highly complex, and without tailored work instructions, the unique products and parts were demanding to assemble correctly. This case study applies to low volume, high variation manufacturing.

<u>**Case 5: Siemens Electronics Factory Erlangen**</u>

Siemens has been awarded a Lighthouse status in five of their factories. For their electronics factory in Erlangen, Germany, which manufactures variable frequency drives, the status was awarded in 2024 (Siemens, 2025). The manufacturing site's market is characterized as high-mix and mid-volume, which calls for adaptability. With the combination of AI, robotics, and digital twins, they have seen 69% improvement in productivity, 40%

decrease in time to market, and 42% decrease in energy consumption. According to Siemens (2025), it was achieved with creating a product digital twin and training AI to visually inspect products, as well as AI-supported logistics and fault detection. New product and production system development are done in virtual environments to ensure that the systems are optimized before entering the real world. Additionally, the digital twin is used for optimizing AGV routes and energy consumption.

The Erlangen factory also managed to design and develop a cutting-edge semiconductor production by utilizing real-time data analysis from critical process steps to understand complex system relationships and increase problem-solving capabilities.



Picture 6. Siemens Experience Center digital twin, Lisbon (Siemens, 2025).

The Siemens Experience Center digital twin in Lisbon utilizes real-time IoT data to create collaboration environments and virtual tours (Picture 6). Additionally, they utilize the twin for automation and production plant optimization. (Siemens, 2025)

Case 6: Konecranes Smart Factory (technology partner Process Genius Oy)

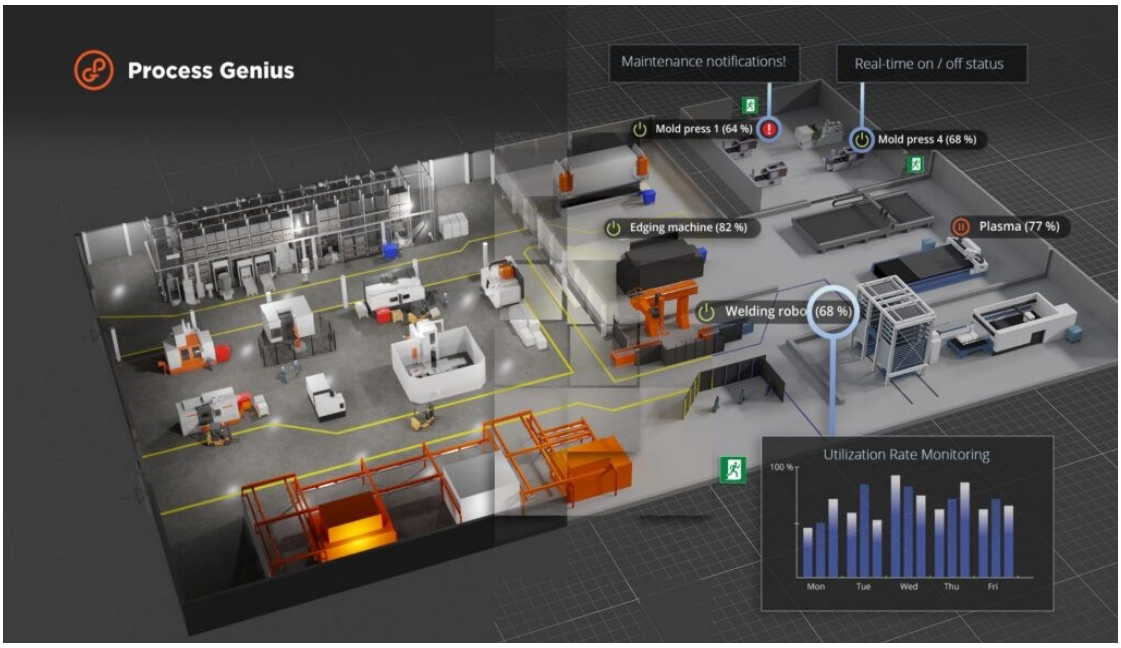
Konecranes smart factory implemented a digital twin together with Process Genius to achieve a smart factory. A three-dimensional twin of the factory was created to monitor real-time factory operations, while safety, business, production, and maintenance are also listed as features and benefits of the system. The data is collected from software, hardware, and sensors, and the system enables monitoring from web browser of mobile device.

Picture 7. Konecranes Smart Factory (Konecranes, 2025).

The benefits of the system are listed as management based on facts, enhanced machine utilization rate, and quick reactions to disruptions. According to Konecranes, their production efficiency increased by 30% in 6 months. Process genius has also implemented a similar solution in other industrial companies, resulting in 30% less hours wasted and 20% reduction in energy consumption. (Eurometalli, 2021; Process Genius, n.d.)

Vendor 1: Process Genius

Genius Core 3D Digital Twin by Process Genius is a SaaS platform which utilizes OPC-UA protocol, and integrates with ERP, OEE, SCADA, and MES, and Azure AD. They offer monthly pricing plans from 950€/month, which includes for example 2D or 3D models, application with notifications, and a multi-level view. They offer a platform which is ready to use after a three-step setup: uploading factory models, connecting IoT data, and training. They provide ready API access, but they do not mention models, sensor data, and data management systems, which should probably be prepared by the customer. Picture 7 presents an example of their smart factory solution. The view includes a 3D layout, and for example maintenance notifications and machine utilization rate monitoring. (Process Genius, n.d.)



Picture 8. Factory digital twin (Process Genius, n.d.).

Vendor 2: Roima Intelligence

Roima Intelligence designs and delivers products related to software and applications. Their factory digital twin integrates information from different sources like MES, ERP or WMS to create a comprehensive view of the whole production process from design to maintenance. In addition to digital twin platforms, they offer multiple smart factory solutions to support the development of a digital twin, such as HMI/SCADA process optimization, advanced planning and scheduling (APS), product digital thread, and data integration solutions. AVEVA Operations Control is a product which utilizes HMI/SCADA integration and provides real-time visualizations and process monitoring while ensuring cross-platform communication and remote access. Additionally, it analyzes data to find patterns. In practice, the AVEVA solution is a digital twin without real-time control or process simulations; the visualizations are provided in graphs and flowcharts. (Roima Intelligence, 2025)

3 Methodology

The purpose of this research is to gather stakeholder requirements for the case company to aid in digital twin design and development. This research aims to define the operational needs for optimal utilization of digital twin and suggest the optimal scope, fidelity and level of analysis based on those needs. Additionally, the goal is to gather anticipated benefits and expected functionalities of the digital twin to create recommendations for the case company.

3.1 Research design and approach

Research onion (Saunders et al., 2007) explains the appropriate philosophy, approach, strategy, choice of methods, time horizon, and techniques. This research design aligns with the objectives of understanding operational needs and optimizing production processes by applying a digital twin in the business unit.

- Philosophy: interpretivism. Seeking to understand the operational needs based on stakeholder expectations and requirements.
- Approach: deductive. The research is based on literature review about digital twins, operational needs and process optimization. The optimal system is then researched, and the possible benefits are analyzed.
- Strategy: a case study. Approach is case study, because it aims to understand and evaluate the operational and stakeholder needs and assess the potential of a twin solution for this specific company.
- Choice of Methods: qualitative. Semi-structured interviews and focus groups with key stakeholders.
- Time horizons: cross-sectional. Current needs and potential benefits are assessed.

Data collection and analysis are done with literature review and using semi-structured interviews with focus groups. The qualitative data is analyzed with thematic analysis.

3.2 Data collection methods

The literature review was done with Finna search service by utilizing key words and filtering by publishing year. The literature is primarily focused on peer-reviewed academic journal articles. This extensive literature evaluation supported the creation of interview questions and the analysis of the results. The qualitative data for this thesis was collected by conducting group interviews with seven different focus groups. A semi-structured interview style was selected, in which the interviewees were asked multiple questions related to five different themes. Interviews are seen as a good data collection method for this kind of research because it ensures that the person or people and their objective experiences are brought in the centre of attention (Saunders et al., 2019). This is crucial when designing a new product. The con of interviewing as a data collection method is that it is time consuming. The questions were open ended, for example: “What are your expectations?” and “What do you want to see and why?”. The purpose was to gather user requirements in the form of “As a process engineer, I want to…”

3.3 Planning and executing the interviews

The aim of the interview was that the results would provide an answer for the main research question: What are the operational needs for a digital twin solution. The interview questions were slightly adapted for different focus groups based in their specific role and expertise, while remaining within the same thematic categories. The operations are divided into seven different subcategories: process and quality, automation, planning, supervising, logistics, project and ramp-up management, and management (Table 3). There were 11 interview questions for each group, which provided 77 answers in total. The interviewees were chosen based on their current role and involvement in the production environment. As all interviewees were native Finnish speakers, the interviews were conducted in Finnish to ensure that the participants felt comfortable answering the questions and no information would be lost. The interviewees were contacted via Teams meeting invitation which included the purpose and the topic of the interviews, and every meeting was recorded and transcribed with Microsoft Teams transcribe -application.

Table 3. Interview information.

Interview group	Area of expertise	No. of participants	Length (h)
Group 1	Process and Quality	5	0:54
Group 2	Automation	5	1:37
Group 3	Production planning	5	1:28
Group 4	Production supervising	6	1:37
Group 5	Logistics	6	1:16
Group 6	Project management and product ramp-up	3	
Group 7	Management	6	1:18

3.4 Data analysis techniques

Microsoft Teams transcribe -feature and recordings of the meetings aided in the raw material collection, which was then sorted into data that is relevant to the research and research questions. The relevant information was found by revising the interview questions, previous studies about this topic and theoretical framework (Saunders et al., 2019). After the relevant information was found, it was analysed with thematic analysis to identify key themes in operational needs.

Analyzing the interview responses by category was done using the requirements management approach. The desired outcome was a list of categorized stakeholder requirements and anticipated benefits, where common themes and key words can be found. Based on the requirements, the optimal solution was defined to make sure all the necessary information and analytics would be accessible in the digital twin. Additionally, data requirements and design structure were defined.

3.5 Reliability of the research

The quality of the interviews and the validity of results is ensured by automatic transcribing and recording of the interviews. Utilizing academic literature before the interviews for question formation and for analyzing the results, the correct interpretation of answers and their significance is ensured. The challenges of requirement detection were recognized as communication barriers, changing requirements, conflicts of needs, incomplete requirements. They were minimized by using the participant's native language, encouraging informal conversation, and asking follow-up questions.

4 Analysis

This chapter introduces the findings from the interviews. The data is analyzed in five thematic categories for the sake of clarity. First, the initial level of knowledge about digital twin technology is analyzed. Then existing operational challenges, data needs and granularity, and anticipated impact of the new system are presented. The last thematic category gathers any specific system requirements that the interviewees might have.

4.1 Level of knowledge about the topic

- *How do you understand digital twin?*

This question was asked from all the focus groups as the first question to map the knowledge of the topic. This was done to ensure the quality of the following data and detect the participating employees' readiness to utilize a digital twin in their work.

Overall, the interviewed groups have a fine understanding of what a digital twin is, and how it can be used in optimization and development of operations. Group 1 described it as a *"Real-time digital copy of some physical world process or thing"*, which is the general definition of a digital twin. All seven groups described it as a digital copy of a physical system, and five of the seven groups described it as real-time, or being connected to the physical world (Figure 20). All the focus groups mentioned simulation. In groups 1, 3, and 4 there were at least one person who had never heard of the system before. Additionally, digital twin was described as a sandbox or a test environment (group 4) and being similar to test SAP (group 3).

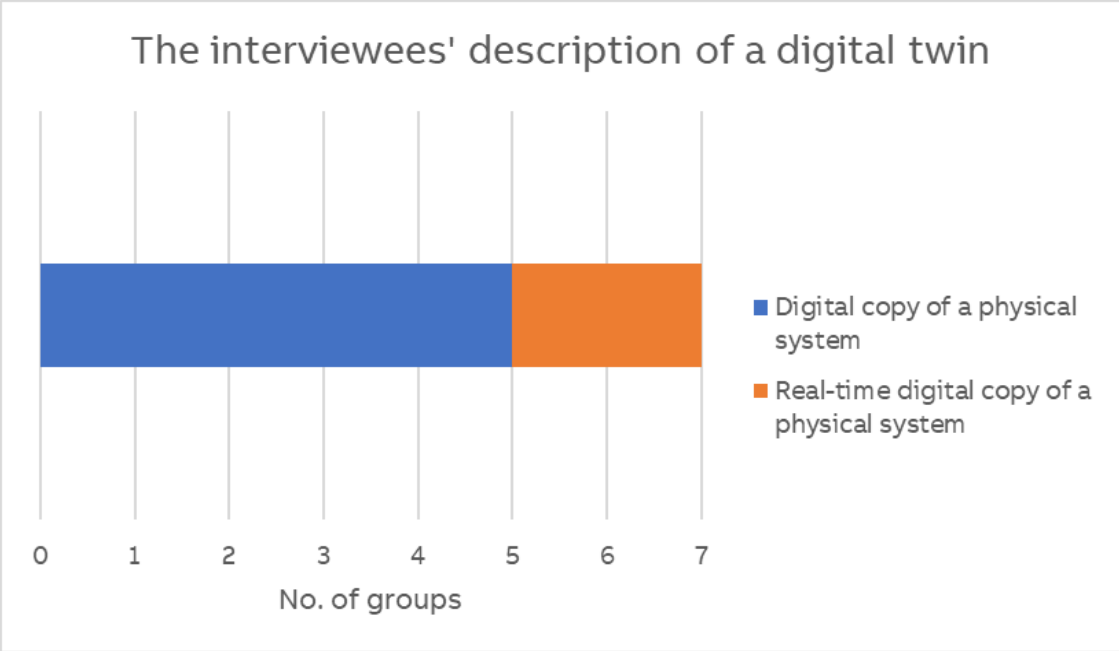


Figure 20. The interviewees' description of a digital twin.

Besides simulation, some other functionalities were discussed by groups 1, 2, 6, and 7. For example, the ability to optimize and simulate (group 1), pre-plan and preview things, and the ability to "view shop floor status from home via VPN" (group 2) were mentioned. Group 6 discussed the system's potential level of control by noting that it could be used with manual control or creating a self-directed twin. Groups 2 and 7 mentioned the possibility for interactivity, and group 7 summarized the desired system usage as modeling, monitoring and optimizing. Additionally, some requirements were already considered by group 2. They questioned what is considered as a real-time system, and what kind of time lag would be optimal; *"Is it a millisecond or half an hour?"* (group 2).

4.2 Existing operational challenges

- *Could you describe the current challenges you face in *your role*?*
- *Can you provide a specific example where real-time data or predictive analysis could have helped you solve a problem in your work?*

This section covers the challenges that the interviewees face in their day-to-day work. Besides asking directly about their challenges, the interviewees were urged to provide examples to provoke more in-depth answers. The current pain points are diverse across all interviewed groups, and there were 15 different challenges mentioned in total (Figure 21). All the interviewed groups mentioned visibility (visibility issues, lack of real time data, or no visibility to root causes) as a challenge. Other frequently mentioned challenges are material issues, which was mentioned by six groups, and communication and coordination issues, which was a current issue in five of the seven groups.

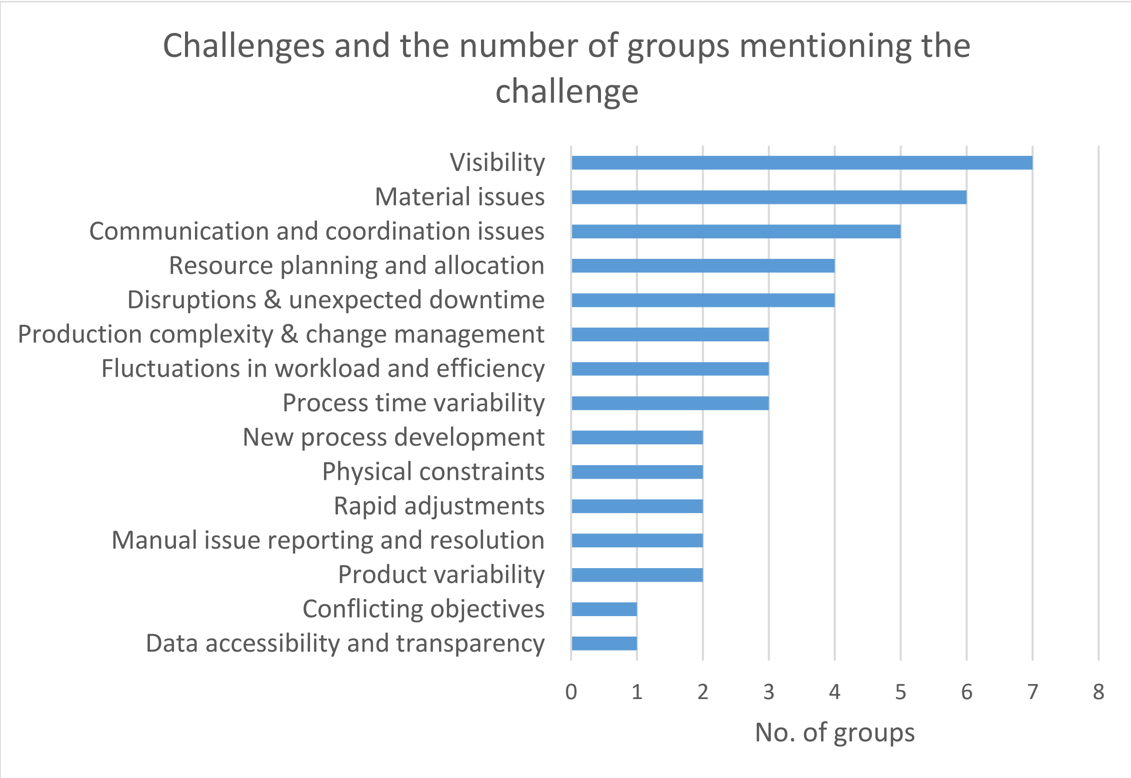


Figure 21. Existing operational challenges.

According to the interviewees, visibility issues originate from unclear reasons behind disruptions, unknown shipment size, unclear material order status, no root cause analysis on missing material, lack of real-time data, difficulty in data synchronizing & extraction, global changes & lack of information regarding the changes, as well as scattered information.

Material issues are described as material availability and restrictions, material requirement being unclear to supporting operations like kitting stations, ordering issues, order status being unclear, manual order behavior which then leads to fluctuations in logistics' workload and missing orders, as well as unclear material status and location. Communication and coordination issues stem from unclear responsibilities, ad-hoc and verbal planning changes, and lack of standardization in disturbance communication. Additionally, information about changes is described as not traveling well between different teams, and customer service communication is described as inefficient.

4.3 Data needs and granularity

- *How do you utilize data or real time information in your work currently, and where are the gaps?*
- *What information would you expect digital twin to monitor *to help you succeed in your role*?*
- *What level of granularity would you need a digital twin to provide to effectively support you in *your role*?*

Data needs and granularity are assessed to detect the most important data and what kind of insights the participants might be lacking. This creates an outline for system scope. Additionally, the answers to this section broadly define the required level of analysis, which is discussed in chapter 4.5 in more detail.

The interviewees express that they currently have a decent amount of data available. They focus on data presentation and processed information rather than requiring more data points. There were 21 data points mentioned across the groups; seven out of the 21 were mentioned by at least three groups and therefore can be considered significant for system design (Figure 22). The top seven data requirements are process flow, material, disruption, workstation, smart tool, tester, and personnel data.

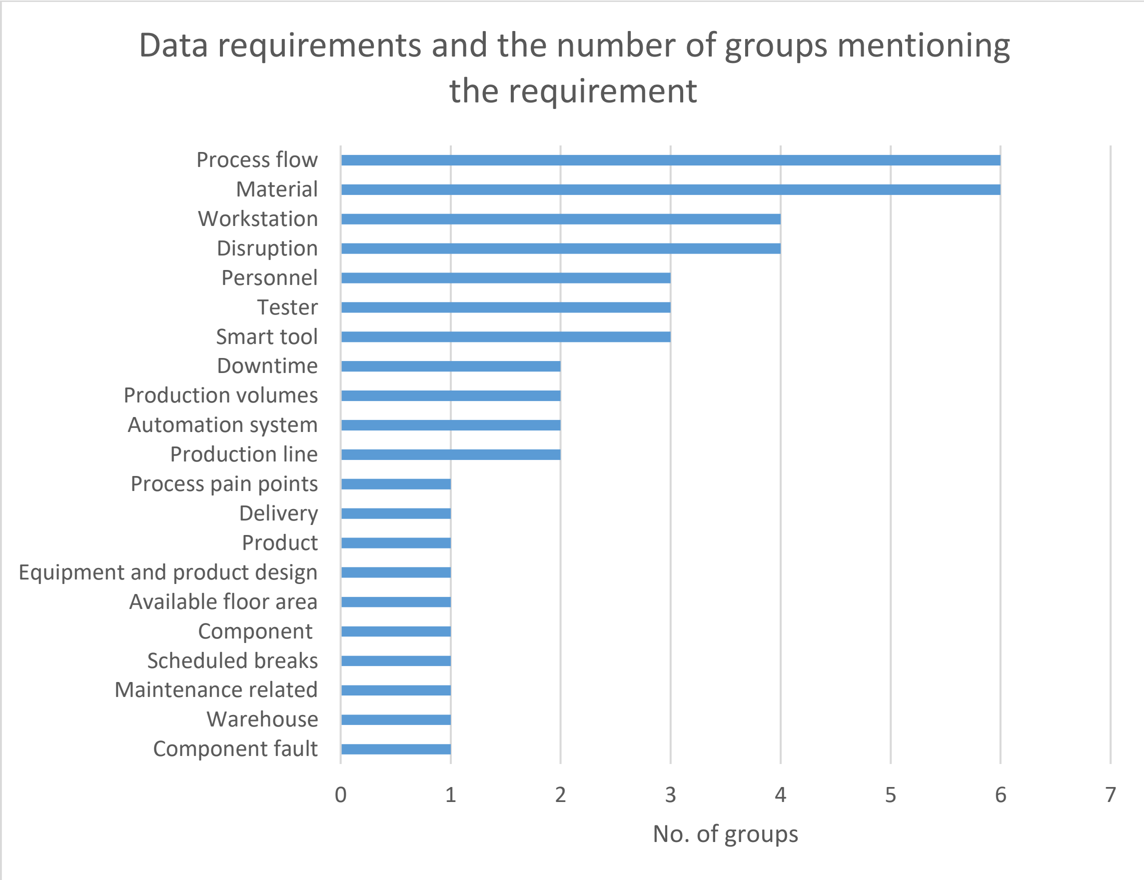


Figure 22. Data requirements and the number of groups mentioning the requirement.

The most mentioned data points were process flow and material related data. All groups except for group 6 mentioned the need to examine process flow data from the new system. This includes operation start and end time, idle time, takt time and efficiency. The participants highlight the need to see especially work stations that are seen as possible bottlenecks: HiPot plus -workstation and packing. The importance of whole process visibility is highlighted by the managers (group 7).

In addition to process flow, other common gap across all groups is material data. In total, there were 10 different material-related data needs mentioned in six groups, and they are material status, order visibility, material location, material availability, allocation visibility, material consumption, material saldo, visibility to material changes, procurement, and supply chain visibility. Group 1 was the only group that did not recognize the need to see material data.

Status was a common theme across all the data requirements, which indicates the need for visibility to the real-time state of workstations, testers, workflow, materials, equipment, automation systems, products, production volume, and personnel (Figure 23). Status was mentioned by groups 1-5 multiple times. Only groups 6 and 7 did not express the need to see the real-time status of the manufacturing environment, suggesting that project & ramp-up management and managers are more focused on the long-time development rather than day-to-day production monitoring. One manager however brought up that customer service often requires the product completion status, to understand whether the product is under work or if it can still be canceled; this is not strictly managerial concern and thus it is left out on the chart. This however reveals the importance of customer service input to requirements detection.

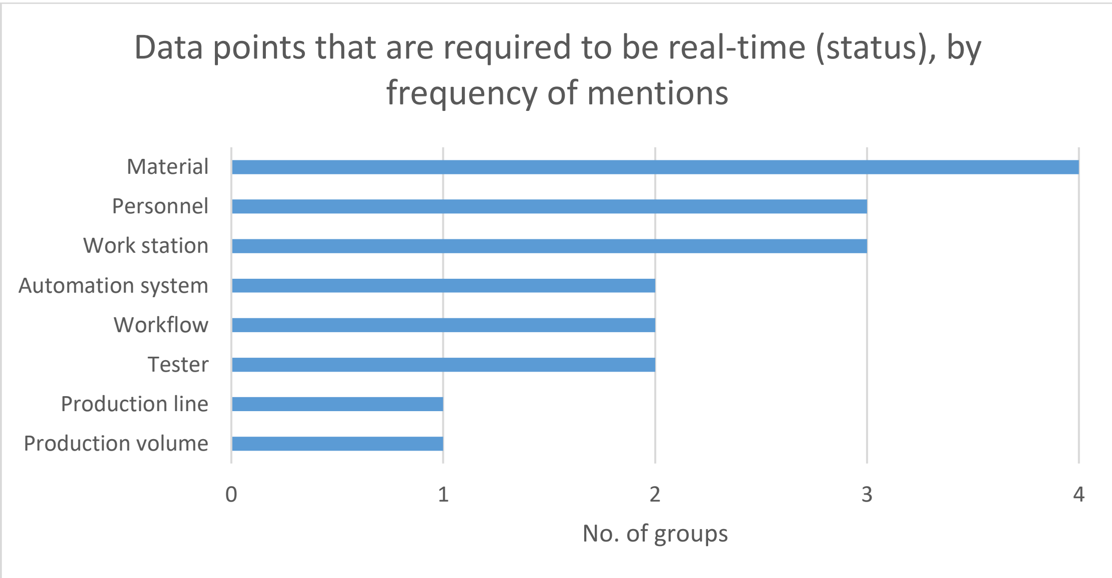


Figure 23. The required real-time information (status) data points of shop floor and the number of groups mentioning the requirement.

All of the participating groups emphasized that there should be a modifiable view of the operations, because data granularity needs vary from process flow and big picture level to component level data depending on user role.

4.4 Anticipated impact of the new system

- *How do you see the digital twin addressing this challenge, if at all?*
- *How could a digital twin help you succeed in *your role*?*
- *How could predictive insights support you in *your role*?*

The participants were asked these three questions to map the expectations each group has for the digital twin. Tables 4 and 5 present the 20 different potential benefits which were mentioned across the interviewed groups, and they are divided into six different themes in Figure 24.

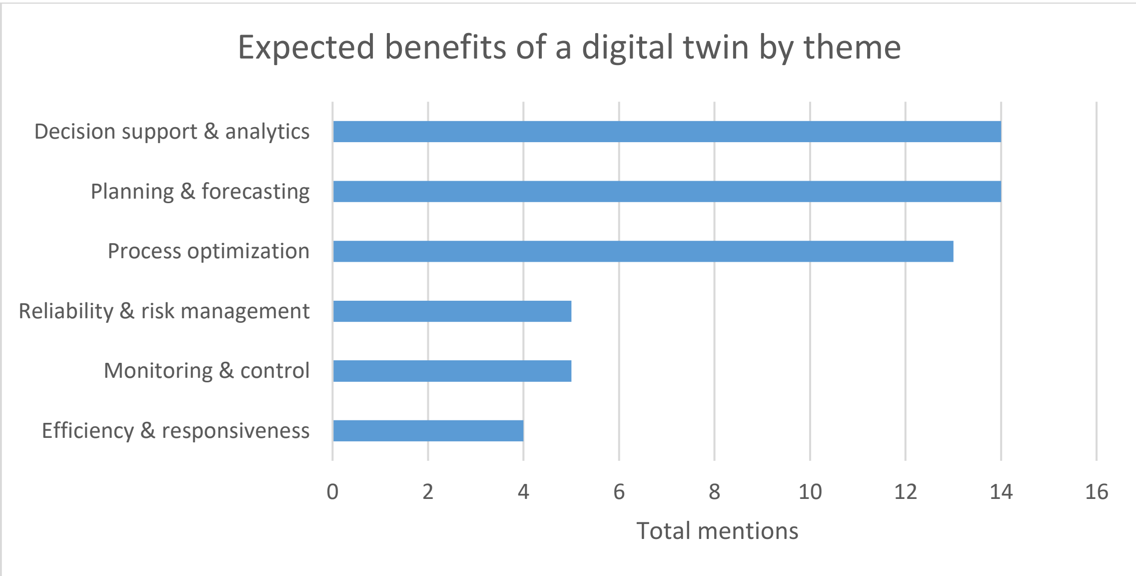


Figure 24. How digital twin is expected to affect the interviewees' work.

Planning and forecasting as well as decision support & analytics related key words were mentioned 14 times while process optimization had 13 mentions in total. Monitoring &

control, reliability & risk management, and efficiency & responsiveness were mentioned four or five times. Specifically, the most discussed benefits were resource planning, scenario planning & simulation, decision support, and bottleneck detection & balancing, which were mentioned by five or more groups (Table 4).

Table 4. The most expected benefits from a digital twin and the number of groups mentioning that benefit.

Planning & forecasting	Decision support & analytics	Process optimization
Resource planning (7) *Scenario planning & simulation (6)* *Future capacity evaluation & development (1)*	*Decision support (5)* *Statistical process control (4)* *Predictive analysis & risk signaling (4)* *Historical data analysis & insights (1)*	*Bottleneck detection & balancing (5)* *Automated scheduling & optimization (4)* *Process and flow optimization (3)* *Variability handling (1)*

Statistical process control, predictive analysis & risk signaling, and automated scheduling & optimization were mentioned in four of the seven groups, and the rest of the expected benefits were identified by less than half of the groups (Tables 4 & 5).

Table 5. Less expected benefits from a digital twin and the number of groups mentioning that benefit.

Monitoring & control	Reliability & risk management	Efficiency & responsiveness
Real-time data visibility (2) *Quality monitoring & control (1)* *Reducing automation downtime (1)* *Automation monitoring (1)*	*Data reliability (2)* *Delivery reliability (2)* *Risk mitigation (1)*	*Own work efficiency (2)* *Faster response to unexpected changes (2)*

However, this does not necessarily imply that these features are less important – the data suggests that different features might be more important to specific focus groups. For example, historical data analysis & insights as well as future capacity evaluation & development are benefits that were mentioned only by managers.

Collaboration

- *How do you anticipate the digital twin impacting collaboration between different teams (production, automation, logistics, management, etc.)?*

Collaboration was not originally mentioned as a benefit of the system, but with a probing question about the effect on cross-team collaboration, the impact was seen as positive (Figure 25). The most expected impact on collaboration is common understanding and common goals, which was mentioned by five groups. Better change transparency and management was emphasized by three groups, and they expressed that the current transparency level is not ideal. The interviewees also stated that the system would result in less communication issues. Other benefits were seen in better disturbance and resource coordination, and resource sharing.

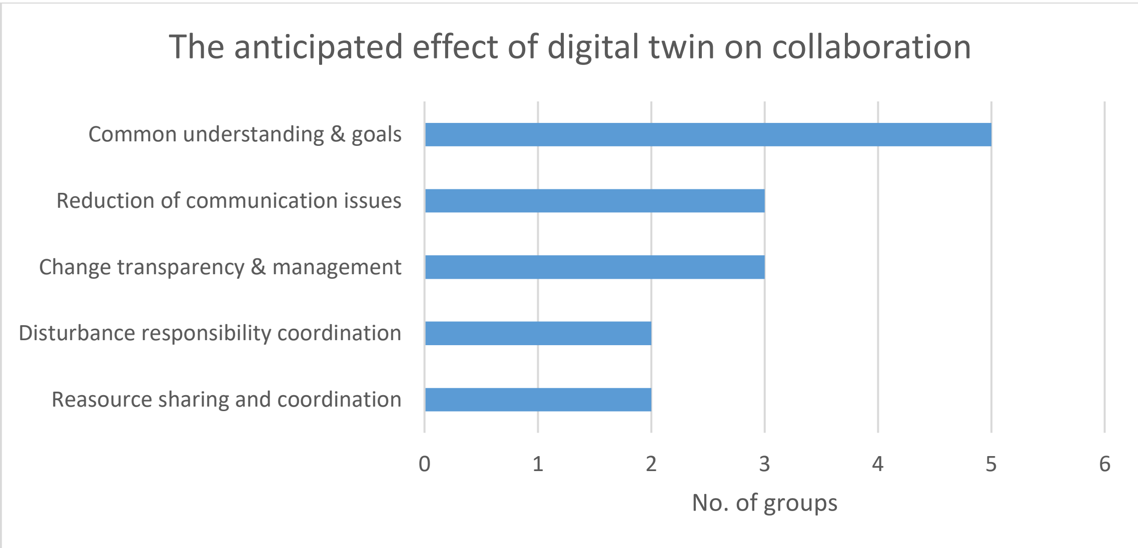


Figure 25. How the participating groups expect digital twin to affect collaboration.

4.5 User requirements

- *Could you elaborate on any specific requirements you would have for a digital twin solution in your role?*

This question was asked to determine if the interviewees have any specific interface preferences, system interactions, feature expectations, usability expectations, or accessibility concerns about the digital twin solution. The answers were divided into system requirements and feature expectations.

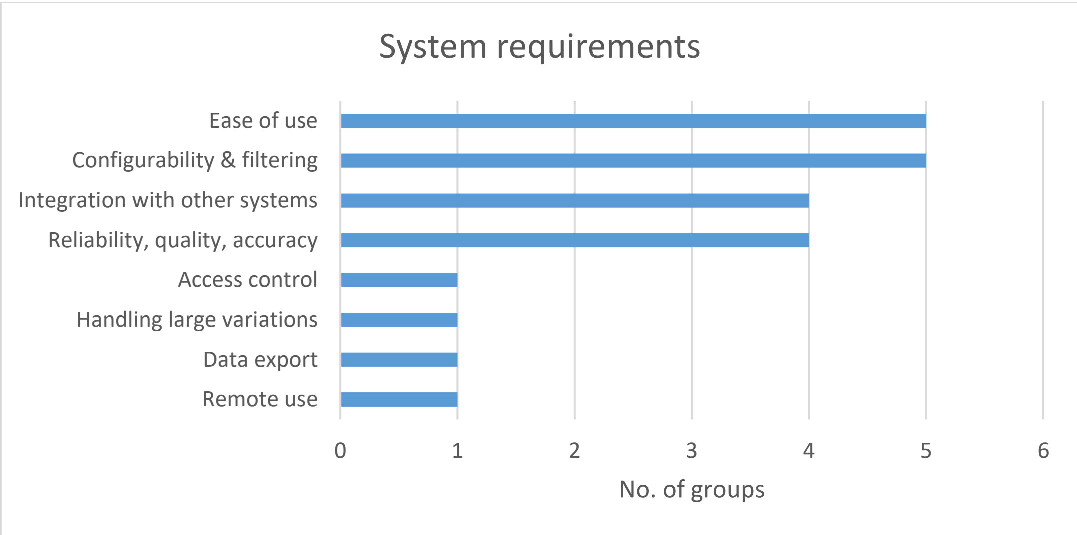


Figure 26. System requirements.

The most common concerns across the interviewed groups were usability and configurability, which were mentioned by 5 groups (Figure 26). Data reliability, quality and accuracy, and integration with other systems were also discussed by most of the interviewees. Other requirements describe a system that should allow remote use and data export. Concerns about access control and the ability to handle large variation of products were raised.

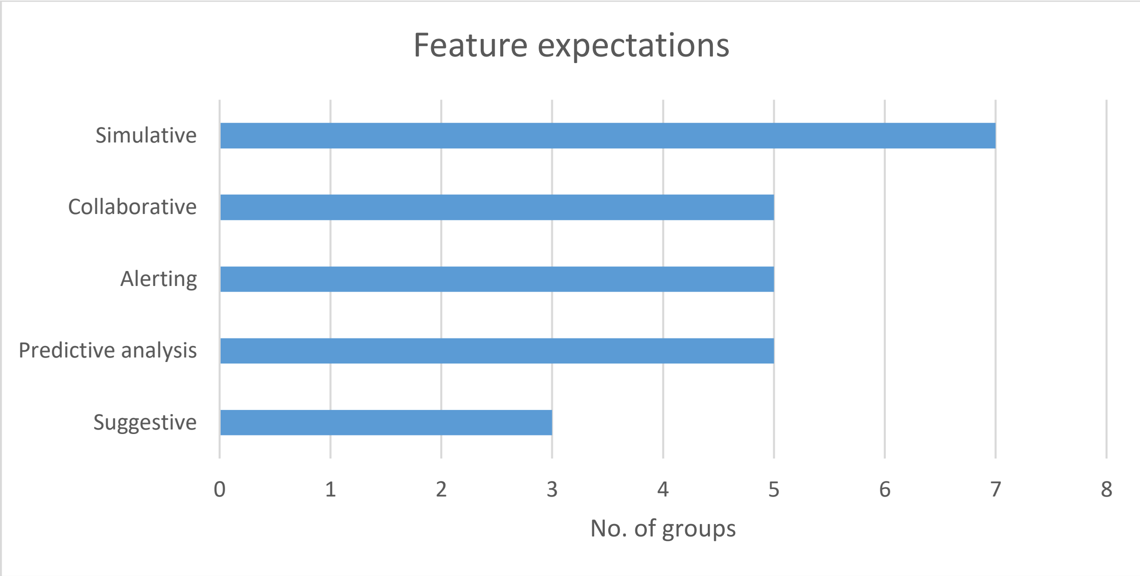


Figure 27. Feature expectations.

There were only five feature expectations mentioned across the participants: They expected simulative, collaborative, alerting, predictive, and suggestive capabilities from the system (Figure 27). All the interviewed groups mention they expect the system to perform simulations. The required simulations were diverse between different roles; some were more focused on effect simulations, for example with absentee or material unavailability effect, while others preferred historical simulations to aid in new system or process development. Five groups mentioned they would benefit from a predictive system. This entails predictive data and predictive analysis, particularly for predictive alerts and problem identification, as well as resourcing suggestions. Additionally, five groups described the need for an alerting or collaborative system. They described the possibility to "leave comments" to the system for example on maintenance related issues or to use it as a communication channel, and the ability to allocate disruption details to the right person. Three groups mention suggestions, which they describe as beneficial especially in resourcing and disturbance handling.

4.6 Analysis conclusion

The key findings from the analysis conclude that the participants have a good understanding and even enthusiasm about the digital twin. The current challenges are diverse, and the data requirements relate to the mentioned challenges. In addition to process flow visibility, real-time material visibility is seen as highly important. Status was mentioned frequently, which suggests that digital twin is an adequate solution to the needs. Impact of the system was anticipated as great and the expected benefits were numerous, especially in decision support and analytics, as well as in planning and forecasting. Usability and configurability are key concerns among the interviewees, and all the groups expect the system to perform simulations, based on both historical data and predictive analysis.

Communication & coordination issues were mentioned as an existing challenge by five of the seven groups, which implies that an optimal solution should include features that support cross-functional transparency. However, the interviewees did not mention enhanced collaboration as an expected benefit at first, but with a probing question, the participants stated that for example common goals, change transparency, and reduction in communication issues are anticipated to be positively impacted. An unexpected finding was the frequency of which resource planning was mentioned. It was stated as a challenge by four groups, and all the interviewed groups mentioned it as an anticipated benefit of the digital twin. By some groups, this was emphasized multiple times. The current literature does not frequently mention this as a main benefit of a digital twin in the industrial setting, but in this case company, the reliance on human labor and the OPF production lines create an environment where the impact of material and human resource fluctuations multiplies the potential disruption time. Thus, improved resource planning can be seen as a valuable benefit in this setting.

The analysis revealed that while generally the current challenges differ between groups and there were 15 different challenges mentioned across all the interviewed groups, there were multiple common challenges related to visibility-, material-, and communication & coordination issues. This was further emphasized by the interviewees' data needs; material and process flow data were mentioned as the top data requirements. Additionally, process optimization was seen as a potential usage for the system. Concern of the new systems ability to handle large variation was raised.

The same type of data can solve multiple different problems for different stakeholders. For example, material data can aid in production planning, resource planning, and space limitations. Access to real-time data (status) supports visibility, which was mentioned as a top challenge. Overall, the findings highlight the need for overall visibility and especially material-related real-time data, suggesting that a digital twin is an ideal fit for the interviewees' needs. These insights will be further discussed in relation to existing literature and practical implications in the following chapter.

5 Results and conclusions

This section presents an overview of the challenges, expected functionalities, required data points, system requirements and feature expectations by connecting key insights to broader implications or theoretical discussions. The purpose is to determine

- What the system must and must not do?
- What operations will this system support?
- What is the required level of detail?

Group-specific needs are briefly discussed to emphasize the differences between each operational function. Additionally, the conclusions aim to define the optimal

- Data integration: model, shadow, or twin
- Fidelity: no of parameters, accuracy, and abstraction
- Level of analysis: descriptive, diagnostic, predictive, prescriptive, or autonomous
- Scope: system, process, product, or component

5.1 Overview

Overall, the interviewed groups have a clear understanding of what the current pain points of each area are, and what they would like the system to provide in order to support day-to-day operations sufficiently. The participants identified multiple different challenges, poor visibility being the most mentioned challenge across all interviewed groups. Other challenges in the top five most pressing issues stem from material issues, communication and coordination issues, disruptions & unexpected downtime, as well as resource planning and allocation (Table 6).

Table 6. Challenges

Top 5 challenges	Other challenges
• Visibility • Material issues • Communication and coordination issues • Disruptions & unexpected downtime • Resource planning and allocation	• Process time variability • Fluctuations in workload and efficiency • Production complexity & change management • Product variability • Manual issue reporting and resolution • Rapid adjustments • Physical constraints • New process development • Data accessibility and transparency • Conflicting objectives

To address the current challenges, the future system must include real-time data and historical data for overall production visibility and adequate root-cause analysis, which allows for better disturbance management. Material issues require multiple real-time data points from supply chain, warehouse, logistics, consumption, and delivery. Additionally, material flow could be visualized for past and future scenario simulation, which aids in new logistics process development and helps determining the right course of action in case of material unavailability. Communication and coordination issues can be addressed with shop-floor visibility and analytics, which help determining the common goals, and a collaborative platform. An alerting and prescriptive system would be optimal in tackling resource planning and allocation difficulties, but simulations of different OPF line resourcing scenarios would be helpful as well.

The most expected benefits are related to planning & forecasting, decision support & analytics, and process optimization (Table 7). Specifically, the participants anticipate the new system will allow better resource- and scenario planning as well as scenario simulation. In addition, bottleneck detection and balancing are functionalities that are expected to optimize processes and decrease idle time of production processes as well as supporting operations. Decision support is expected especially among managers and people who are mitigating process optimization in cases of disturbance or material restrictions, for example production planners and supervisors. Other expected functionalities are seen as less important for the system to satisfy user expectations.

Table 7. Expected benefits

Most expected benefits	Expected benefits
• Planning & forecasting • Decision support & analytics • Process optimization	• Monitoring & control • Reliability & risk management • Efficiency & responsiveness

These functionalities require different levels of analytics. For example, resource planning and decision support are functionalities that require the twin to have optimization capabilities; they need the platform to integrate data from different places, simulate, and create analytics based on the comparison of real-world and simulated operations. In an ideal situation, a prescriptive twin would perform these simulations continuously in the background and provide advice on the optimal changes to the system. Scenario planning & simulation as well as bottleneck detection are possible with a digital shadow or a simulation, and do not require as much advanced analytics. Forecasting requires predictive analytics, which is attainable with large amounts of production data from multiple data sources for a long period of time.

The participants defined that the system must include data from process flow, materials, disruption details, and workstation status. Other data points like smart tool related data, testing data, personnel, production line, automation system, volume and downtime are seen as should-have information, because they were not mentioned by as many focus groups. The rest of the data points are not emphasized as necessary points for the whole operations but might be necessary for specific groups and thus are labeled under the nice-to-have category in Table 8.

Table 8. Data requirements

Must have data	Should have data	Nice to have data
• Process flow data • Material data • Disruption details • Workstation status	• Smart tool • Tester • Personnel • Production line • Automation system • Production volumes • Downtime	• Component fault • Warehouse • Maintenance related • Scheduled breaks • Component • Available floor area • Equipment and product design • Product • Delivery • Process pain points

The data requirements suggest that the system must include high-level data from production process flow and workstation status, enhanced with more detailed data of material status. Especially material order visibility and location are required. Additionally, disruption details are seen as necessary for the quick resolution and prevention of downtime and should be visible and preferably already analyzed information in the digital twin platform. This combination of high-level operational data and detailed granular data requires a multi-level twin development, because it includes different subsystems and data levels. This multi-level approach is scalable and supports the later incorporation of data points that were categorized as should-have data in this research.

The interviewees were asked one question about any possible requirements they might have for the system. The answers were categorized into system requirements (Table 9), which describe how the system should function, and feature expectations (Table 10), which reflect the complexity and desired level of analysis. The features that the digital twin must have include ease of use, configurability and filtering, as well as reliability, quality, and accuracy. Additionally, the system must integrate well with other systems. These features can be seen as non-negotiable. Additional features add value but are secondary to the core needs.

Table 9. System requirements

Most important requirements	Important requirements
• Ease of use • Configurability & filtering • Reliability, quality, accuracy • Integration with other systems	• Remote use • Data export • Handling large variations • Access control

The most important requirements should not be compromised, and similar are discussed in literature and case studies. Additionally, the important requirements like remote use and data export add value to certain stakeholders and should be considered important. Some participants were concerned of the system's ability to handle the large variations of products and having access control, and even though these requirements were mentioned only once, they are vital to the security and adaptability of the system.

All the participants agreed that the system must include simulations. The interviewees emphasized the importance of predictive alerts, predictive problem identification, and predictive resourcing suggestions and described how it would positively affect for example production downtime and idle time. Alerts and a collaborative environment, where for example messages and details are allocated directly to the right person, were mentioned by most of the groups. A suggestive system was mentioned by three of the seven groups, and they stated that it would be extremely useful for example in resourcing and disturbance handling.

Table 10. Feature expectations

Must have features	Should have features	Nice to have features
• Simulations	• Collaborative • Alerting • Predictive analysis	• Suggestive (prescriptive)

5.2 Group-specific needs

Every group has a different perspective depending on their role and objective. The group-specific needs are examined to gain understanding of the most prominent differences in data needs, system requirements, and expected features.

The material related data needs are vastly different for every group (Figure 28). Group 1 was the only group that did not mention data needs related to material, and it was most important for group 3 (production planning). They require information about material status, orders, location, availability, allocation visibility, and balance (saldo). Group 5 (logistics) require material status, order visibility, and location, but also consumption information.

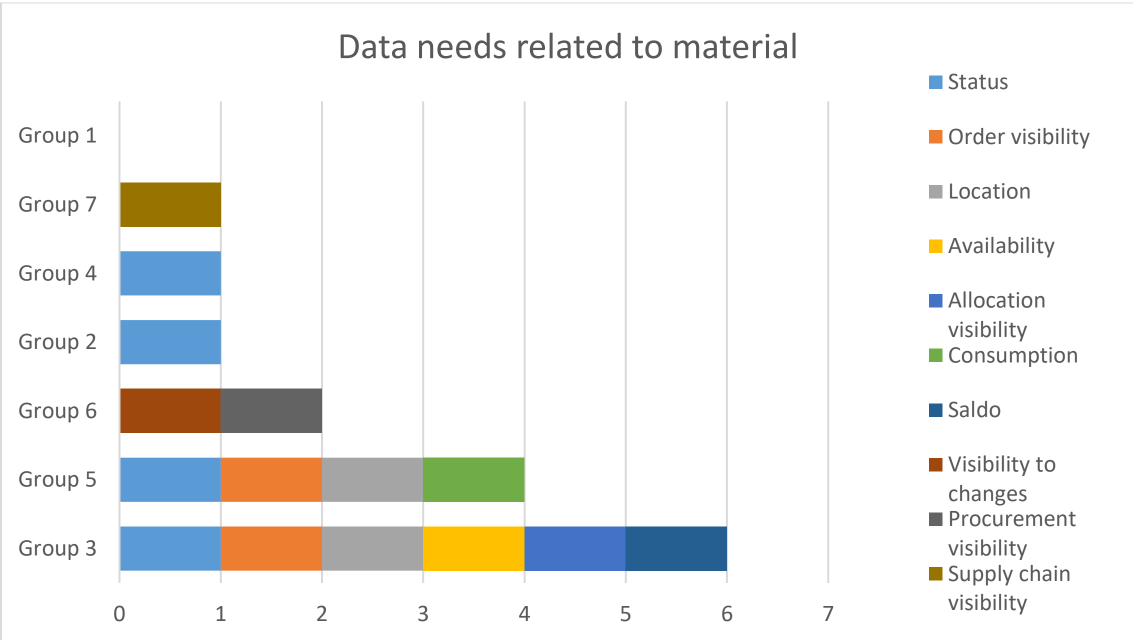


Figure 28. Material data requirements by group.

Process and quality engineers' data needs can be summarized as predictive insights, smart tools, and streamlined, processed information rather than raw data. This group values automation and insights over raw data. They require information that helps predict problems and optimize processes proactively. They emphasized the importance of integration with other systems, and expected high level of analytics, such as suggestions, which translates to prescriptive analytics, predictive analytics, and alerts.

Especially automation engineers and specialists focused on data requirements, and they stated that data must be in common format, all in the same place, and presented well. Additionally, they suggested avoiding too much data to keep the system functional. Distinctive data needs for this group were detailed diagnostics, system-wide visibility, simulation environment for robots, and a modifiable system. This group required granular system data and fault logs for diagnostics and optimizing performance, but also a mix of real-time and historical data to track maintenance and bottlenecks. Some of their requirements were remote use, access control, integration with other systems, and configurability as well as a collaborative platform.

Production planning focuses on material data availability. This group prioritizes real-time alerts, resource allocation suggestions, and scenario simulations for disruptions, absentee effects, and material shortages. This group needs real-time and predictive data to improve decision-making in a rapidly changing production environment. They focus on scenario simulations and material availability data. Some of the features that they expect from the digital twin include suggestions, predictions, and alerts, for example material or resource unavailability alerts.

Production supervisors had similar data needs to production planning. This is because both groups operate in close proximity to production, thus experiencing the same types of problems. The supervisors need to see what is currently happening everywhere on the shopfloor. They requested a predictive and alerting system, which would automatically allocate disruption details to the right person. Additionally, their system requirements were firm: reliability, quality, accuracy, and usability.

For the logistics team, material data, like consumption and visibility to orders would result in more timely and optimized operations. They required visibility to production to anticipate material and resourcing needs proactively. This group discussed about the benefits of logistics simulations to aid in planning and predictive disruption management. In addition to simulation, some other expected features are predictive, alerting, and collaborative capabilities.

Project and product ramp-up management would mostly benefit from a product digital twin. They often need details from different teams, for example procurement details, design details, material changes, product changes, and equipment design details. Especially ramp-up management requires screw level details to products. System requirements include handling large variations, which is vital for the development of new products and production systems. Additionally, they expected the system to provide ready, pre-configured simulations, rather than requiring users to create them independently.

Management stated that they would mostly benefit from big picture visibility. Additionally, they highlighted the need for a scalable system and minimum viable product design, which this research aims to define. They expected the optimal system to provide advanced analytics like predictions of future states and suggestions on future actions. Integration with other systems was emphasized as highly important by the managers.

5.3 Conclusions

The challenges and the most mentioned data needs together suggest that the system must provide real-time data from production process flow and workstation status, and more detailed data of material status and disruption details. The minimum viable product should primarily support production process flow by providing a real-time view of the shop floor processes, especially of the production lines and cell workstations but also including material. Top priority system requirements are ease of use and configurability. Integration with current systems as well as reliability, quality, and accuracy of the system are seen as highly important. Other requirements like access control, the ability to handle large variations, data export, and remote use are not discussed as frequently. This could be due to the expectation that these qualities are self-evident. However, in the literature they are considered highly important and should not be overlooked in system design.

5.3.1 Answers to research questions

RQ1: What are the operational needs for a digital twin solution?

To address the current challenges, the system must create visibility across the operations by providing real-time data about the main production processes and material flow from procuring to delivery. The system must have planning and forecasting capabilities in order to satisfy user expectations. Additionally, the digital twin should have decision support and analytics capabilities along with process optimization tools. These expectations can be met with a predictive and prescriptive twin. Additionally, the literature suggests that a twin that allows for real-time control can help minimizing unexpected downtime.

RQ2: What is the optimal process coverage and granularity for the digital twin based on the operational needs?

An optimal solution would be a multi-level, modular, and scalable process digital twin. A minimum viable product includes high-level process flow data and workstation status, enhanced with more granular disruption details and material flow data. Material data, especially status, order visibility and location, is seen as vital, which is supported by the fact that material issues are prominent in the current operational challenges. The optimal granularity is different for every user, and the view should be modifiable. Product related requirements or current challenges were mentioned only a few times; component and product design information were not mentioned as frequently, and they were described to aid in customer service visibility, new product ramp-up, and faulty material visibility. However, these needs were not stressed as significant in current challenge detection, which suggests that product or component twins can be left out of the primary scope.

RQ3: How can the digital twin be used in optimizing the manufacturing operations?

Three different levels of required fidelity were recognized: descriptive, predictive and prescriptive. A descriptive twin would be useful for scenario planning, and a real-time descriptive twin enhances overall visibility, aid in communication providing a unified understanding of the past, higher level resource planning for future plans, scenario planning, bottleneck detection and balancing, and production flow visualization. However, this solution without advanced analytics would not fully satisfy user needs. A platform which provides predictive and prescriptive analytics is able to tackle the most prominent operational challenges and enhance efficiency in material flow optimization, communication and coordination, and disruption management. It is able to predict future issues and propose solutions which can prevent issues before they occur. Ultimately, a prescriptive twin can be utilized as a strategic decision-making tool.

Predictive analytics is required in predictive alerts, problem identification, and resourcing suggestions, and it can aid in material issues, unexpected downtime and disruption managing, and even decision support. Prescriptive system is most beneficial in resource planning, decision support, and automated scheduling & optimization, as well as process and flow optimization. However, literature suggests that an autonomous twin can create substantial benefits for example in automation control. This type of real-time control is useful when there is a need for autonomous actions, but this was not a feature that any of the interviewees mentioned.

Finally, based on this research, the optimal digital twin for ABB Drive Product's operations is not in fact categorized as a digital twin, but a digital shadow. This is supported by the fact that all of the data requirements and most expected functionalities can be addressed with a solution that collects, analyzes, and presents real-time data and simulations, but does not include automatic data flow back to the physical system. Additionally, a digital shadow is able to aid in all of the current challenges that were mentioned in the interviews. This suggests that it is a beneficial tool for this manufacturing environment.

6 Discussion and recommendations

This chapter discusses the key findings of the research on the optimal digital twin solution for ABB Drive Products' operations and presents potential improvements on three KPIs (downtime, idle time, overall efficiency) as an example. Additionally, recommendations for the case company are provided, along with suggestions for future research.

6.1 Potential benefits for Drive Products' operations

Benefits from the literature review align with current problems and the benefits that the interviewees expected from the system. Similar systems have provided benefits for example in downtime reduction and efficiency gains. In this case company, operational data suggests that besides downtime, there are inefficiencies in process balancing and bottlenecks. The potential effect can be roughly estimated from similar production environments that have implemented digital twins in their operations.

As stated in the chapter 2.5, Siemens drives factory in Germany accomplished a 69% enhancement in productivity with AI, digital twin, and automation together. For Konecranes, a Process Genius solution increased production efficiency by 30% in 6 months. Process Genius has also implemented a similar solution in other industrial companies, resulting in 30% less hours wasted and 20% reduction in energy consumption. Additionally, a production line digital twin case study by Mendi (2022) showed a downtime reduction from approximately 200 minutes per month to around 20; an 87% reduction in a six-month follow-up period. This was accomplished with an alerting decision support system.

6.2 Recommendations

The research found that most operational challenges originate from visibility issues, material issues, and communication & coordination challenges. Therefore, it is recommended that the company focuses on improving these issues and integrates a system that focuses on real-time data, simulative capabilities, and providing a shared platform for common understanding. Additionally, the minimum viable product should include process flow-, material-, and disruption data, as well as workstation status. To enable the upgrade from minimum viable product to a higher fidelity solution, the system should be designed with scalability in mind. Ultimately, a multi-fidelity digital twin for optimization provides the possibility to utilize simplified, low-fidelity simulations for real-time process monitoring and high-fidelity simulations periodically to update the twin for long-term accuracy. It effectively balances cost and precision.

When deciding between internal development and external acquisition, it is recommended that the company considers their core competency, long-term vision, and budget and timeline constraints (Daneshgar et al., 2013). This ensures that the development aligns with the company's strategy, and that time-to-value is understood. Additionally, there are benefits and limitations to both options. Internal development ensures customization, control, flexibility, and intellectual property ownership, while it may delay operational benefits and require significant investment in software development and talent acquisition. When done externally, faster implementation and industry expertise are provided along with technical support and updates. However, this results in less customization, possible integration challenges, and vendor lock-in. (Shahzad et al., 2017)

Since this technology is not within the core competency of ABB Drive Products, external acquisition of the foundational platform is recommended. A chosen technology partner should allow the expansion of the twin as the operations change and expand. Additionally, they should provide different levels of maturity, from minimum viable product and predictive analytics to autonomous actions if needed. Customized solutions and data security are essential requirements. This approach ensures that the initial implementation is fast. Additionally, the minimum viable product is created with the optimal technical support, while the internal developers gain the experience to further develop new modules, customize, and maintain the digital twin. However, internal resources are needed for the development and maintenance, and the development should be led internally to ensure a smooth transition from external platform provider to internal upkeep. While several companies advertise platforms with the required features, a comprehensive evaluation of their suitability for our specific needs requires direct engagement and further discussion with the companies.

Whether the case company develops the twin internally or acquires it from an external company, either as a tailored solution or as a subscription, the IT and OT architecture needs to be established. Existing operational technology should be either updated or integrated with emerging technologies, and future technology investments should consider compatibility with modern communication protocols such as OPC-UA.

6.3 Future research

This research has provided the background of digital twin technology, what are the different types and levels of digital twins, and the core concepts of data gathering and utilization. The implementation process, required technologies, and different software were also researched. However, digital twin technology is still in its development phase, and thus research, standardization, tools, and literature are immature, and future research should focus on finding the optimal tools and technologies for this specific company. This requires continuous research of new trends and technologies to stay on top of this field. Additionally, the cost of the chosen technology or system should be carefully optimized together with required fidelity and expected benefits. This further assesses system viability, ensuring that the benefits will outweigh the high initial cost.

The user needs have been determined in relation to data, system fidelity, and expected functionalities. Optimal data source, format, and collection method must be subsequently defined. This entails the determination of current available data, data quality, if there are any gaps, and if more data should be gathered for example through sensors. Additionally, non-functional and technical requirements must be defined - this includes information like how many simultaneous users should the system support and how will the system provide it. Collaboration with the IT department and digital transformation team ensures that data management is executed properly.

This is preliminary research. Considering the importance of a comprehensive view of the costs and benefits, as well as the optimal fidelity and the synergic nature of the twin, it is important to activate all stakeholders in the process and product lifecycle before defining the system scope completely, especially if something other than a minimum viable solution is developed. User needs were defined with seven stakeholder groups, who will be using the system in the shop floor manufacturing operations. This has involved process and quality, automation, production planning and management, logistics, project and product ramp-up, and management. Further research should determine if other related departments or roles would benefit from the system as well; for example, product management and sales have their own requirements, which should be discovered. A twin that includes product, process, and system could create synergic advantages. These operations' requirements should also be verified against strategic goals. A digital twin has the capability to reduce energy consumption and provide a more sustainable way of production, but these needs did not surface in the detection of requirements.

References

Abanda, F. H., Jian, N., Adukpo, S., Tuhaise, V. V., & Manjia, M. B. (2025). Digital twin for product versus project lifecycles' development in manufacturing and construction industries. *Journal of Intelligent Manufacturing*, *36*(2), 801–831. https://doi.org/10.1007/s10845-023-02301-2

Agrawal, A., Thiel, R., Jain, P., Singh, V., & Fischer, M. (2023). Digital Twin: Where do humans fit in? *Automation in Construction*, *148*, 104749. https://doi.org/10.1016/j.autcon.2023.104749

Alimam, H., Mazzuto, G., Tozzi, N., Emanuele Ciarapica, F., & Bevilacqua, M. (2023). The resurrection of digital triplet: A cognitive pillar of human-machine integration at the dawn of industry 5.0. *Journal of King Saud University - Computer and Information Sciences*, *35*(10), 101846. https://doi.org/10.1016/j.jksuci.2023.101846

Anaya, V., Alberti, E., & Scivoletto, G. (2024). A Manufacturing Digital Twin Framework. In J. Soldatos (Ed.), *Artificial Intelligence in Manufacturing* (pp. 181–193). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-46452-2_10

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

Boyes, H., Hallaq, B., Cunningham, J., & Watson, T. (2018). The industrial internet of things (IIoT): An analysis framework. *Computers in Industry*, *101*, 1–12. https://doi.org/10.1016/j.compind.2018.04.015

Standish Group. (2020). *Chaos Report 2020*. Standish Group.

Chen, C., Miao, Q., Kumar, S., & Wu, T. (2023). Privacy-preserving authentication scheme for digital twin-enabled autonomous vehicle environments. *Transactions on Emerging Telecommunications Technologies*, *34*(11), e4751. https://doi.org/10.1002/ett.4751

Conti, P., Guo, M., Manzoni, A., & Hesthaven, J. S. (2023). Multi-fidelity surrogate modeling using long short-term memory networks. *Computer Methods in Applied Mechanics and Engineering*, *404*, 115811. https://doi.org/10.1016/j.cma.2022.115811

Corallo, A., Del Vecchio, V. D., Lezzi, M., & Morciano, P. (2021). Shop Floor Digital Twin in Smart Manufacturing: A Systematic Literature Review. *Sustainability*, *13*(23), 12987. https://doi.org/10.3390/su132312987

Daase, C., Haertel, C., Nahhas, A., Volk, M., Steigerwald, H., Ramesohl, A., Schneider, B., Zeier, A., & Turowski, K. (2023). Following the Digital Thread – A Cloud-Based Observation. *Procedia Computer Science*, *217*, 1867–1876. https://doi.org/10.1016/j.procs.2022.12.387

Daneshgar, F., Low, G. C., & Worasinchai, L. (2013). An investigation of 'build vs. buy' decision for software acquisition by small to medium enterprises. *Information and software technology,* *55*(10), 1741-1750. https://doi.org/10.1016/j.infsof.2013.03.009

Desislavov, R., Martínez-Plumed, F., & Hernández-Orallo, J. (2023). Trends in AI inference energy consumption: Beyond the performance-vs-parameter laws of deep learning. *Sustainable Computing: Informatics and Systems*, *38*, 100857. https://doi.org/10.1016/j.suscom.2023.100857

Dihan, Md. S., Akash, A. I., Tasneem, Z., Das, P., Das, S. K., Islam, Md. R., Islam, Md. M., Badal, F. R., Ali, Md. F., Ahamed, Md. H., Abhi, S. H., Sarker, S. K., & Hasan, Md. M. (2024). Digital twin: Data exploration, architecture, implementation and future. *Heliyon*, *10*(5), e26503. https://doi.org/10.1016/j.heliyon.2024.e26503

Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2018). *Fundamentals of Business Process Management*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-56509-4

Ehie, I. C., & Chilton, M. A. (2020). Understanding the influence of IT/OT Convergence on the adoption of Internet of Things (IoT) in manufacturing organizations: An empirical investigation. *Computers in Industry*, *115*, 103166. https://doi.org/10.1016/j.compind.2019.103166

Grieves, M. (2016). *Origins of the Digital Twin Concept*. Unpublished. https://doi.org/10.13140/RG.2.2.26367.61609

Grieves, M. (2022). Intelligent digital twins and the development and management of complex systems. *Digital Twin*, *2*, 8. https://doi.org/10.12688/digitaltwin.17574.1

Gupta, S., & Starr, M. (2014). *Production and Operations Management Systems* (0 ed.). CRC Press. https://doi.org/10.1201/b16470

Ibrahim, M., Rjabtšikov, V., & Gilbert, R. (2023). Overview of Digital Twin Platforms for EV Applications. *Sensors*, *23*(3), 1414. https://doi.org/10.3390/s23031414

Kober, C., Algan, B. N., Fette, M., & Wulfsberg, J. P. (2023). Relations of Digital Twin Fidelity and Benefits: A Design-to-Value Approach. *Procedia CIRP*, *119*, 809–815. https://doi.org/10.1016/j.procir.2023.03.126

Kober, C., Fette, M., & Wulfsberg, J. P. (2023). A Method for Calculating Optimum Digital Twin Fidelity. *Procedia CIRP*, *120*, 1155–1160. https://doi.org/10.1016/j.procir.2023.09.141

Kuehn, W. (2018). Digital twins for decision making in complex production and logistic enterprises. *International Journal of Design & Nature and Ecodynamics*, *13*(3), 260–271. https://doi.org/10.2495/DNE-V13-N3-260-271

Lattanzi, L., Raffaeli, R., Peruzzini, M., & Pellicciari, M. (2021). Digital twin for smart manufacturing: A review of concepts towards a practical industrial implementation. *International Journal of Computer Integrated Manufacturing*, *34*(6), 567–597. https://doi.org/10.1080/0951192X.2021.1911003

Lehner, C., Padovano, A., Zehetner, C., & Hackenberg, G. (2024). Digital twin and digital thread within the product lifecycle management. *Procedia Computer Science*, *232*, 2875–2886. https://doi.org/10.1016/j.procs.2024.02.104

Lehner, D., Pfeiffer, J., Tinsel, E.-F., Strljic, M. M., Sint, S., Vierhauser, M., Wortmann, A., & Wimmer, M. (2022). Digital Twin Platforms: Requirements, Capabilities, and Future Prospects. *IEEE Software*, *39*(2), 53–61. https://doi.org/10.1109/MS.2021.3133795

Li, W., Zhou, H., Lu, Z., & Kamarthi, S. (2024). Navigating the Evolution of Digital Twins Research through Keyword Co-Occurence Network Analysis. *Sensors*, *24*(4), 1202. https://doi.org/10.3390/s24041202

Liu, J., Liu, J., Zhuang, C., Liu, Z., & Miao, T. (2021). Construction method of shop-floor digital twin based on MBSE. *Journal of Manufacturing Systems*, *60*, 93–118. https://doi.org/10.1016/j.jmsy.2021.05.004

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

Lorenz, R., Kraus, M., Wolf, H., Feuerriegel, S., & Netland, T. H. (2024). Selecting advanced analytics in manufacturing: A decision support model. *Production Planning & Control*, *35*(7), 711–724. https://doi.org/10.1080/09537287.2022.2126951

Maier, J. F., Eckert, C. M., & John Clarkson, P. (2017). Model granularity in engineering design – concepts and framework. *Design Science*, *3*, e1. https://doi.org/10.1017/dsj.2016.16

McClenaghan, A., Gopsill, J., Ballantyne, R., & Hicks, B. (2023). Cost Benefit Analysis for Digital Twin Model Selection at the Time of Investment. *Procedia CIRP*, *120*, 1197–1202. https://doi.org/10.1016/j.procir.2023.09.148

Mendi, A. F. (2022). A Digital Twin Case Study on Automotive Production Line. *Sensors*, *22*(18), 6963. https://doi.org/10.3390/s22186963

Mourtzis, D. (2020). Simulation in the design and operation of manufacturing systems: State of the art and new trends. *International Journal of Production Research*, *58*(7), 1927–1949. https://doi.org/10.1080/00207543.2019.1636321

Nguyen, T. H., Le, X. C., & Vu, T. H. L. (2022). An Extended Technology-Organization-Environment (TOE) Framework for Online Retailing Utilization in Digital Transformation: Empirical Evidence from Vietnam. *Journal of Open Innovation: Technology, Market, and Complexity*, *8*(4), 200. https://doi.org/10.3390/joitmc8040200

Nuseibeh, B., & Easterbrook, S. (2000). Requirements engineering: A roadmap. *Proceedings of the Conference on The Future of Software Engineering*, 35–46. https://doi.org/10.1145/336512.336523

Onaji, I., Tiwari, D., Soulatiantork, P., Song, B., & Tiwari, A. (2022). Digital twin in manufacturing: Conceptual framework and case studies. *International Journal of Computer Integrated Manufacturing*, *35*(8), 831–858. https://doi.org/10.1080/0951192X.2022.2027014

Pandey, D., Suman, U., & Ramani, A. K. (2010). An Effective Requirement Engineering Process Model for Software Development and Requirements Management. *2010 International Conference on Advances in Recent Technologies in Communication and Computing*, 287–291. https://doi.org/10.1109/ARTCom.2010.24

Perno, M., Hvam, L., & Haug, A. (2022). Implementation of digital twins in the process industry: A systematic literature review of enablers and barriers. *Computers in Industry*, *134*, 103558. https://doi.org/10.1016/j.compind.2021.103558

Phalak, P., Tomba, E., Jehoulet, P., Kapitan-Gnimdu, A., Soladana, P. M., Vagaggini, L., Brochier, M., Stevens, B., Peel, T., Strodiot, L., & Dessoy, S. (2023). Digital Twin Implementation for Manufacturing of Adjuvants. *Processes*, *11*(6), 1717. https://doi.org/10.3390/pr11061717

Pronost, G., Mayer, F., Camargo, M., & Dupont, L. (2023). Digital Twins along the product lifecycle: A systematic literature review of applications in manufacturing. *Digital Twin*, *3*, 3. https://doi.org/10.12688/digitaltwin.17807.1

Resman, M., Protner, J., Simic, M., & Herakovic, N. (2021). A Five-Step Approach to Planning Data-Driven Digital Twins for Discrete Manufacturing Systems. *Applied Sciences*, *11*(8), 3639. https://doi.org/10.3390/app11083639

Rivera-Pinto, A., Kildal, J., & Lazkano, E. (2024). Toward Programming a Collaborative Robot by Interacting with Its Digital Twin in a Mixed Reality Environment. *International Journal of Human–Computer Interaction*, *40*(17), 4745–4757. https://doi.org/10.1080/10447318.2023.2221599

Sahoo, S., Kumar, S., Abedin, M. Z., Lim, W. M., & Jakhar, S. K. (2023). Deep learning applications in manufacturing operations: A review of trends and ways forward. *Journal of Enterprise Information Management*, *36*(1), 221–251. https://doi.org/10.1108/JEIM-01-2022-0025

Saunders, M., Thornhill, A., & Lewis, P. (2019). *Research methods for business students* (Eigth Edition). Pearson.

Shahzad, B., Abdullatif, A. M., Ikram, N., & Mashkoor, A. (2017). Build Software or Buy: A Study on Developing Large Scale Software. *IEEE access, 5*, 24262-24274. https://doi.org/10.1109/ACCESS.2017.2762729

Skorupińska, E., Hitka, M., & Sydor, M. (2024). Surveying Quality Management Methodologies in Wooden Furniture Production. *Systems*, *12*(2), 51. https://doi.org/10.3390/systems12020051

Thompson, N. C., Greenewald, K., Lee, K., & Manso, G. F. (2020). *The Computational Limits of Deep Learning* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2007.05558

VanDerHorn, E., & Mahadevan, S. (2021). Digital Twin: Generalization, characterization and implementation. *Decision Support Systems*, *145*, 113524. https://doi.org/10.1016/j.dss.2021.113524

Xia, T., Sun, H., Ding, Y., Han, D., Qin, W., Seidelmann, J., & Xi, L. (2025). Digital twin-based real-time energy optimization method for production line considering fault disturbances. *Journal of Intelligent Manufacturing*, *36*(1), 569–593. https://doi.org/10.1007/s10845-023-02219-9

Xie, J., & Wan, J. (2023). Digital Twin Four-Dimension Fusion Modeling Method Design and Application to the Discrete Manufacturing Line. *Big Data and Cognitive Computing*, *7*(2), 89. https://doi.org/10.3390/bdcc7020089

Yi, Y., Yan, Y., Liu, X., Ni, Z., Feng, J., & Liu, J. (2021). Digital twin-based smart assembly process design and application framework for complex products and its case study. *Journal of Manufacturing Systems*, *58*, 94–107. https://doi.org/10.1016/j.jmsy.2020.04.013

Yuan, B., Zhang, J., Lyu, A., Wu, J., Wang, Z., Yang, M., Liu, K., Mou, M., & Cui, P. (2024). Emergence and Causality in Complex Systems: A Survey of Causal Emergence and Related Quantitative Studies. *Entropy*, *26*(2), 108. https://doi.org/10.3390/e26020108

Zhang, Z., Guan, Z., Gong, Y., Luo, D., & Yue, L. (2022). Improved multi-fidelity simulation-based optimisation: Application in a digital twin shop floor. *International Journal of Production Research*, *60*(3), 1016–1035. https://doi.org/10.1080/00207543.2020.1849846

Zheng, X., Lu, J., & Kiritsis, D. (2022). The emergence of cognitive digital twin: Vision, challenges and opportunities. *International Journal of Production Research*, *60*(24), 7610–7632. https://doi.org/10.1080/00207543.2021.2014591

Zhuang, C., Miao, T., Liu, J., & Xiong, H. (2021). The connotation of digital twin, and the construction and application method of shop-floor digital twin. *Robotics and Computer-Integrated Manufacturing*, *68*(102075). https://doi.org/10.1016/j.rcim.2020.102075

Appendices

Appendix 1. Interview questions

Level of knowledge about the topic

- How do you understand digital twin?

Existing operational challenges

- Could you describe the current challenges you face in *your role*?
- Can you provide a specific example where real-time data or predictive analysis could have helped you solve a problem in your work?

Anticipated impact of the new system

- How do you see the digital twin addressing this challenge, if at all?
- How could a digital twin help you succeed in *your role*?
- How could predictive insights support you in *your role*?
- How do you anticipate the digital twin impacting collaboration between different teams (production, automation, logistics, management, etc.)?

Data needs and granularity

- How do you utilize data or real time information in your work currently, and where are the gaps?
- What information would you expect digital twin to monitor *to help you succeed in your role*?
- What level of granularity would you need a digital twin to provide to effectively support you in *your role*?

User requirements

- Could you elaborate on any specific requirements you would have for a digital twin solution in your role?