

# **Automatic Generation and Validation of Simulation Based Digital Twins**

## **A Data-driven Framework and Case Study Application**

---

**Author:** Daniel Fischer

**Supervisor:** Prof. Christian Schwede

**Program:** Research Master Data Science

**Institution:** Hochschule Bielefeld (HSBI)

**Submission Date:** March 11, 2025

# Contents

<b>List of Abbreviations</b>	<b>6</b>
<b>1 Introduction</b>	<b>7</b>
1.1 Initial Situation . . . . .	7
1.2 Problem . . . . .	10
1.3 Objective of the Thesis . . . . .	11
1.4 Structure of the Thesis and Methodological Approach	11
<b>2 Introduction</b>	<b>12</b>
2.1 Initial Situation . . . . .	12
2.2 Problem . . . . .	13
2.3 Objective of the Thesis . . . . .	14
2.4 Structure of the Thesis und Methodological Approach	15
<b>3 Theoretical Foundations and State of the Art</b>	<b>17</b>
3.1 Digital Twin: Definition and Concepts . . . . .	18
3.1.1 Types of Digital Twins	18
3.1.2 Data-Driven Digital Twin (DT)	18
3.1.3 Rationale from the Paper	18
3.2 Data-Driven Modeling: Principles and Methods . . .	18
3.3 Material Flow Planning and Simulation . . . . .	18
3.3.1 Fundamental Concepts from Christian's Book	18
3.3.2 Processes and Resources	18
3.3.3 Insights from Pinedo's Book	18
3.3.4 Production Planning and Control	18

3.4	Validation and Verification in Simulation . . . . .	18
3.4.1	Standard Formats and Data Sources for Simulation Models	18
3.4.2	Validation Methods from Christian’s Book	18
3.4.3	Key Figures and KPIs	18
3.4.4	Presentation of My Standard Format, <i>Pro Conta</i> , versus Other Formats	18
3.4.5	Trace Analysis	18
3.4.6	Challenges in Validation	18
 4	 <b>State of Research</b>	 <b>—19</b>
4.1	Existing Approaches for the Validation and Verifica- tion of Digital Twins . . . . .	19
4.2	Automatic Model Generation and Its Reasons/Chal- lenges . . . . .	19
4.2.1	Always Up-to-Date and Online Validation (Twin no longer rep- resents reality)	19
4.2.2	High Complexity Requires Automated Procedures	19
4.2.3	Framework Promotes Standardization and Transparency	19
4.2.4	Scalability of Automated Approaches	19
4.2.5	Manual Validations May Contain Bias	19
4.2.6	Cost Savings	19
4.3	Limitations of Current Standard Formats and Data Structures . . . . .	19
4.4	Gaps and Open Questions in Research . . . . .	19
 5	 <b>Methodology and Framework Development</b>	 <b>—20</b>
5.1	Requirements Analysis . . . . .	20
5.2	Conceptualization . . . . .	20
5.2.1	Description of the Framework for Automatic Model Generation	20
5.2.2	Verification of the Model: Structured Pipeline and Validation Goals	20
5.2.3	Data-Driven Validation: Event Data and KPIs for System Eval- uation	20
5.2.4	Definition of Validation Scenarios (e.g., Masking, Blackbox Sce- narios)	20

5.2.5	Automatic Evaluation Criteria for Model Validity (e.g., similarity to $R^2$ , Throughput Times, Efficiency Metrics)	20
5.2.6	Handling of Missing or Limited Data	20
<b>6</b>	<b>Implementation of the Framework</b>	<b>21</b>
6.1	Technical Implementation of Model Generation . . .	21
6.1.1	UML Diagram	21
6.1.2	Integration into Existing Systems	21
6.2	Automatic Validation of the Generated Model . . . .	21
6.3	Use of External Information Sources for Model Verification . . . . .	21
6.4	Interfaces to Material Flow Systems and Data Collection . . . . .	21
<b>7</b>	<b>Case Study: Validation of a Digital Twin in the Production System</b>	<b>22</b>
7.1	Description of the Production System and Available Data . . . . .	22
7.1.1	Description of the IoT Factory	22
7.1.2	Data Preparation and Analysis	22
7.2	Construction of the Digital Twin for the Scenario . .	22
7.3	Conducting the Validation Experiments . . . . .	22
7.4	Results and Interpretation: Limits and Errors of the Model . . . . .	22
7.5	Limits and Falsifiability of the Model Based on Real Data . . . . .	22
<b>8</b>	<b>Discussion of the Results</b>	<b>23</b>
8.1	Critical Reflection on the Framework Development .	23
8.2	Limitations of Automated Validation . . . . .	23
8.3	Significance and Robustness of the Developed Methods . . . . .	23
8.3.1	Comparison with Manual Modeling	23

8.4	Implications for Research and Practice . . . . .	23
<b>9</b>	<b>Conclusion and Outlook</b> —	<b>24</b>
9.1	Summary of the Key Results . . . . .	24
9.2	Answering the Research Questions . . . . .	24
9.3	Outlook: Possible Further Development of the Frame- work . . . . .	24
9.4	Recommendations for Practical Application . . . . .	24

# List of Abbreviations

**CPS** Cyber-Physical System. 1

**DES** Discrete-Event Simulation. 1

**DM** Digital Model. 1

**DMFS** Discrete Material Flow Systems. 1

**DS** Digital Shadow. 1

**DT** Digital Twin. 1

**IoT** Internet of Things. 1

**V&V** Verification and Validation. 1

# Chapter 1

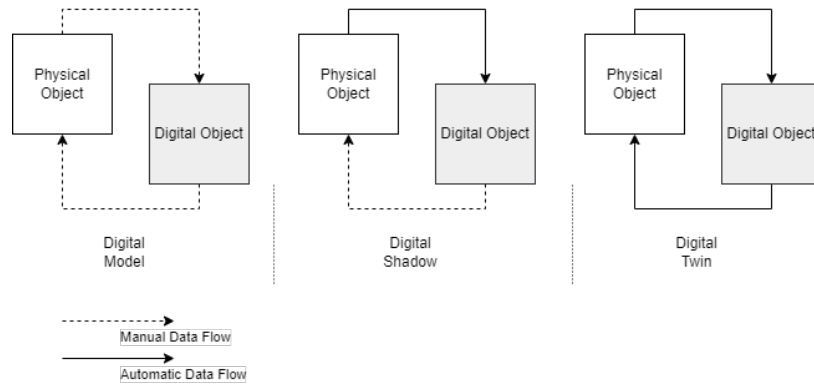
## Introduction

### 1.1 Initial Situation

Digital Twins are a key technology at the front of the fourth industrial revolution, coined Industry 4.0. Latter term is characterized by the convergence of cyber-physical systems (CPS), the Internet of Things (IoT), and cloud computing to create smart factories with the goal of automation and efficiency (Oztemel & Gursev, 2020). Companies pursue the goal to remain competitive by adapting innovative technologies that promise enhanced productivity and reduced operational costs. One such technology promising this is the digital twin (DT). It can be defined as a virtual representation of physical assets enabling real-time monitoring and optimization (Tao et al., 2018). The twin is an enabling technology, bridging the connection between the two entities with a bi-directional data flow to exchange information and to influence the behaviour of the physical asset (Grieves, 2014). This technology in Industry 4.0 connects the physical and digital worlds through real-time data integration, simulation, and optimization (Judijanto et al., 2024).

Although this field is rapidly evolving, there has not been found a common ground about the definition of DT. The term was first introduced by Michael Grieves in 2002, defining it as a digital representation of a physical object or system (Grieves, 2014). However, the concept has evolved since then, encompassing a broader range of applications and technologies. Going back through the literature, there are three words for describing similar characteristics of DT: Digital Model (DM), Digital Shadow (DS) and Digital Twin (DT), see Figure 1.1 (Jones et al., 2020; Zhang et al., 2021).

The DM represents the most basic form. It contains manual data connections be-



**Figure 1.1:** Comparison of Digital Shadow (DS), Digital Model (DM) and Digital Twin (DT) as presented by Kritzinger. This distinction is crucial for understanding validation requirements across different digital representation types.

tween physical and digital entities. These connections can be temporarily shifted or even disconnected. There is no control of the digital object over the physical entity. It rather is a simple or complex model *describing* the modelled object. It can not make decisions by itself to influence the physical object. The reason lies in the potential outdated data the digital part possesses or in the fact that it does not contain logic to control the data flow back to the physical part by itself. The control and the obligation to interpret the results is completely in the hands of the modeller. The DS is a more advanced version of the DM. It is a digital representation of the physical object, which is continuously updated with real-time data. The DS can be used for monitoring, analysis and simulation purposes. It can predict the future state of the physical object based on the current state and historical data. However, the DS is not able to influence the physical object. The control is, similar to the DM, still in the hands of the modeller. A DS is mainly used for simulation and in the literature often confusingly classified as a DT. The DT is the most advanced version of the triplet. It is a digital representation of the physical object, which is also continuously updated with real-time data. The DT can be used for monitoring, analysis, and *control* purposes. It can predict the future state of the physical object based on the current state and historical data. The DT can also influence the physical object by sending control signals to it. The control is partially or completely in the hands of the DT. The DT thus *can* serve more purpose than modelling or simulating the physical object. It may serve as an autonomic system, updating itself or by help of the modeller (Kritzinger et al., 2018).

Digital twins are applied across various sectors, including manufacturing, defense, automotive, and recycling. This thesis focuses on manufacturing, particularly discrete material flow systems (DMFS). These systems process discrete



objects (parts) that move along transportation routes or conveyor lines—either at regular or irregular intervals—integrating both production and logistics operations (Arnold & Furmans, 2005; Schwede & Fischer, 2024). A core simplification in their modeling is the abstraction of material flow as a sequence of discrete events, similar with the principles of discrete-event simulation (DES) (Kovacs & Kostal, 2016; Robinson, 2014). DES is particularly well-suited for analyzing complex systems where state changes occur at discrete points in time, such as arrivals, departures, and processing steps (Robinson, 2014).

In hindsight, DM played a crucial role in the design, planning, and control of DMFS, primarily through applications such as material flow simulations, logistic assistance systems, and digital factory implementations (Thiede et al., 2013). However, advancements in both DS and DT have enabled a revolution from isolated, usecase-specific models toward complete digital models that span the entire lifecycle of DMFS (Abdoune et al., 2023). This transition is largely driven by the increasing demand for predictive capabilities by stakeholders and automated decision support in manufacturing systems, reflecting the cornerstones of Industry 4.0 (Frank et al., 2019).

In practice, the automated data transfer between the digital model and the physical system is of secondary importance for DMFS management. Unlike in time-critical applications, human decision-makers remain an integral part of the control loop, ensuring that real-time automation is not always necessary. Consequently, for this thesis, digital simulations and digital twins will be treated as equivalent concepts.

Beyond merely replicating the current state and managing historical data, digital twins serve a crucial function in predicting system behavior and evaluating potential modifications. The widespread adoption of DES within digital twins highlights the central role of simulation-based digital twins (SBDT) in DMFS (Lugaresi & Matta, 2021). As Schwede and Fischer, 2024 emphasize, SBDTs provide decision support for optimizing cost and performance in highly competitive manufacturing environments. While current SBDTs are primarily developed and updated manually by domain experts, emerging research explores how machine learning (ML) can enhance predictive accuracy and automate model updates by automatically learning model characteristics, reducing costs and development time.

Thus, the progression from digital models to simulation-based digital twins reflects an ongoing shift towards data-driven, predictive, and increasingly automated representations of DMFS, ensuring more informed decision-making through-

out the system lifecycle.

## 1.2 Problem

Despite the immense transformative potential of DT, their implementation comes with significant challenges. The creation and maintenance of accurate Digital Twins require investment in technology and domain knowledge. This investment yields no return if the resulting model fails to accurately represent the modelled entity or delivers incorrect results. As industries integrate DT into their production processes, establishing trust becomes fundamental (Arrieta et al., 2020; Trauer et al., 2022). To gain widespread acceptance, the technology must demonstrate accuracy, transparency, and cost-efficiency (Shao et al., 2023; Wright & Davidson, 2020). Without these qualities, organizations will likely fall back to familiar methods, potentially building resistance to technological advancement. The European Union

One of the key barriers to adoption lies in validation and verification (VV). Ensuring the validity, reliability, and accuracy of a Digital Twin is critical, yet traditional VV approaches rely heavily on manual expert involvement and case-specific reference values. This creates inefficiencies, particularly in the context of automated Digital Twin generation, where such manual processes conflict with the goal of reducing development effort.

For discrete material flow systems, these challenges are even more pronounced due to their dynamic nature and inherent stochastic elements. When Digital Twins for these systems are generated automatically, conventional validation approaches become particularly problematic, as they negate much of the efficiency gained through automation. This creates an apparent paradox: automated Digital Twin generation reduces initial development effort but potentially increases subsequent validation complexity.

This thesis addresses this paradox by proposing that object-centric event logs—the same data structures often used to generate these twins—can serve as the foundation for an automated, use case-independent validation and verification framework. Such an approach would maintain the efficiency benefits of automated generation while ensuring the resulting Digital Twins meet necessary quality standards.

### **1.3 Objective of the Thesis**

### **1.4 Structure of the Thesis and Methodological Approach**

# Chapter 2

## Introduction

Diese Arbeit befasst sich mit der automatisierten Verifikation und Validierung von automatisiert generierten, simulationsbasierten digitalen Zwillingen für diskrete Materialflusssysteme. Im Zuge der Digitalisierung industrieller Prozesse und der vierten industriellen Revolution (Industrie 4.0) gewinnen digitale Zwillinge zunehmend an Bedeutung. Die vorliegende Einleitung skizziert zunächst die aktuelle Situation und den daraus resultierenden Bedarf, formuliert das zentrale Problem sowie die Zielsetzung der Arbeit und erläutert abschließend den Aufbau sowie die methodische Vorgehensweise.

### 2.1 Initial Situation

Im Rahmen der vierten industriellen Revolution findet eine verstärkte Integration von Informations- und Kommunikationstechnologien in Produktions- und Logistikprozesse statt. Digitale Zwillinge, also virtuelle Repliken physischer Systeme, ermöglichen nicht nur die Überwachung von Anlagen, sondern auch deren dynamische Simulation und Optimierung. Besonders in diskreten Materialflusssystemen – wie sie in der Automobilindustrie, der Logistik oder in hochautomatisierten Fertigungsprozessen vorkommen – erweist sich die präzise Modellierung von Materialflüssen als Schlüsselfaktor für Effizienz und Wettbewerbsfähigkeit (**Grieves2014**).

Die rasante Entwicklung im Bereich der Sensorik und des Internet of Things (IoT) liefert kontinuierlich große Datenmengen, die in digitalen Zwillingen genutzt werden können, um den aktuellen Zustand von Produktionssystemen nahezu in Echtzeit abzubilden. Traditionelle Simulationsansätze, die oft auf statischen Modellen basieren, werden durch die dynamische Natur digitaler Zwillinge ergänzt,

die durch kontinuierliche Datenintegration eine verbesserte Prozesssteuerung und prädiktive Instandhaltung ermöglichen (**Tao2018**). Dabei stellen digitale Zwillinge eine Schnittstelle zwischen der realen und der virtuellen Welt dar und bieten somit die Möglichkeit, komplexe industrielle Abläufe zu überwachen, zu simulieren und zu optimieren.

Die zunehmende Digitalisierung und Vernetzung führen jedoch auch zu neuen Herausforderungen. Die Komplexität moderner Systeme sowie die enorme Datenflut erfordern innovative Ansätze zur Modellierung und kontinuierlichen Aktualisierung. Hierbei steht insbesondere die automatisierte Generierung von Simulationsmodellen im Vordergrund, die durch datengetriebene Verfahren unterstützt wird. Dieser Paradigmenwechsel fordert nicht nur traditionelle Simulationen heraus, sondern eröffnet auch neue Perspektiven für die industrielle Praxis, indem er die Grundlage für eine optimierte Entscheidungsfindung schafft (**Uhlemann2017**).

## 2.2 Problem

Trotz des hohen Potenzials digitaler Zwillinge treten in der Praxis zahlreiche Herausforderungen auf. Die automatisierte Erstellung simulationsbasierter Modelle birgt das Risiko, dass wesentliche Systemkomponenten oder Interaktionsmuster unzureichend erfasst werden. Dies führt zu Diskrepanzen zwischen dem digitalen Modell und dem realen System, was im schlimmsten Fall die Grundlage für fehlerhafte Simulationen und falsche betriebliche Entscheidungen bildet. Traditionelle Verifikations- und Validierungsprozesse (V&V) basieren häufig auf manuellen Prüfungen, die angesichts der Komplexität und der Datenmengen in modernen Produktionssystemen nicht mehr effizient durchführbar sind (**Kritzinger2018**).

Ein weiteres Problemfeld stellt die Integration von Echtzeitdaten in die Simulationsmodelle dar. Während historische Daten die Basis statischer Modelle bilden, erfordert die dynamische Anpassung digitaler Zwillinge die kontinuierliche Auswertung und Synchronisation aktueller Betriebsdaten. Dies führt zu Herausforderungen hinsichtlich der Datenkonsistenz, -qualität und -verarbeitung. Zudem fehlt es oft an standardisierten Bewertungsmetriken, die eine objektive Beurteilung der Modellgüte ermöglichen. Diese Probleme machen deutlich, dass herkömmliche V&V-Methoden nicht ohne Weiteres auf automatisiert generierte, simulationsbasierte digitale Zwillinge übertragbar sind.

Insbesondere in diskreten Materialflusssystemen, in denen zahlreiche Prozess-

variablen und Interaktionskomponenten eine Rolle spielen, erfordert die Sicherstellung der Modellgenauigkeit neue Ansätze. Es besteht daher dringender Forschungsbedarf, um datengetriebene und automatisierte V&V-Methoden zu entwickeln, die den spezifischen Anforderungen solcher Systeme gerecht werden (**Kritzinger2018; Uhlemann2017**).

## 2.3 Objective of the Thesis

Das Hauptziel dieser Arbeit ist die Entwicklung eines datengetriebenen Frameworks zur automatisierten Verifikation und Validierung von digital generierten, simulationsbasierten digitalen Zwillingen. Der Fokus liegt dabei auf diskreten Materialflusssystemen, da diese aufgrund ihrer Komplexität und Dynamik besondere Herausforderungen hinsichtlich der Modellgenauigkeit und Prozesssynchronisation aufweisen.

### Motivation und Relevanz

Die Motivation dieser Forschungsarbeit ergibt sich aus mehreren zentralen Aspekten:

- **Effizienzsteigerung:** Durch den Einsatz automatisierter V&V-Methoden können Modellabweichungen frühzeitig erkannt und behoben werden, was zu einer signifikanten Reduktion von Fehlerquellen und Optimierungspotenzialen in Produktionsprozessen führt.
- **Digitalisierung und Vernetzung:** Mit der zunehmenden Integration von IoT und Sensorik in Produktionsanlagen steigt die Verfügbarkeit von Echtzeitdaten, die in digitalen Zwillingen verarbeitet werden können. Ein zuverlässiges V&V-Framework stellt sicher, dass diese Daten adäquat in die Modelle einfließen.
- **Wettbewerbsvorteile:** Unternehmen, die in der Lage sind, ihre digitalen Modelle präzise und effizient zu validieren, sichern sich langfristig Wettbewerbsvorteile in einem zunehmend globalisierten Marktumfeld (**Grieves2014**).

### Forschungsfragen und Hypothesen

Die Arbeit fokussiert sich auf folgende Forschungsfragen:

- Wie können automatisierte Prozesse zur Verifikation und Validierung von digital generierten, simulationsbasierten digitalen Zwillingen effizient implementiert werden?

- Welche datengetriebenen Ansätze eignen sich am besten, um Abweichungen zwischen simuliertem Verhalten und realen Betriebsdaten in diskreten Materialflusssystemen zu identifizieren?
- Inwieweit verbessert das entwickelte Framework die Qualität und Zuverlässigkeit der digitalen Zwillinge im Vergleich zu traditionellen V&V-Methoden?

Es wird die Hypothese aufgestellt, dass ein integriertes, datengetriebenes V&V-Framework signifikante Verbesserungen in der Modellgenauigkeit und Effizienz der Validierungsprozesse bewirken kann. Durch den Einsatz moderner Datenanalysen und maschineller Lernverfahren sollen systematische Abweichungen frühzeitig erkannt und Korrekturmaßnahmen automatisiert eingeleitet werden (Tao2018).

## 2.4 Structure of the Thesis und Methodological Approach

Der Aufbau dieser Arbeit gliedert sich in mehrere Kapitel, die im Folgenden kurz vorgestellt werden:

- **Kapitel ??:** Umfassende Literaturrecherche. In diesem Kapitel werden die theoretischen Grundlagen digitaler Zwillinge, bestehende Simulationansätze und aktuelle V&V-Methoden analysiert.
- **Kapitel ??:** Entwicklung des datengetriebenen Frameworks. Hier werden die konzeptionellen und architektonischen Entscheidungen, die Auswahl der Algorithmen und die Implementierungsdetails dargestellt.
- **Kapitel ??:** Fallstudie. Anhand eines konkreten Beispiels aus dem Bereich der diskreten Materialflusssysteme wird das entwickelte Framework empirisch validiert.
- **Kapitel ??:** Evaluation. Die Ergebnisse der Fallstudie werden hinsichtlich Modellgenauigkeit, Prozesszeiten und Fehlerhäufigkeiten quantitativ und qualitativ ausgewertet.
- **Kapitel 9:** Zusammenfassung und Ausblick. Abschließend werden die wesentlichen Erkenntnisse zusammengefasst, Limitationen diskutiert und Perspektiven für zukünftige Forschungsarbeiten aufgezeigt.

Methodologisch basiert die Arbeit auf einem iterativen Entwicklungs- und Validierungsansatz, der sowohl qualitative als auch quantitative Methoden integriert.

Zunächst erfolgt eine detaillierte Literaturrecherche, um den aktuellen Stand der Technik zu erfassen und bestehende Forschungslücken zu identifizieren. Aufbauend auf diesen Erkenntnissen wird das Framework in modularer Form entwickelt. Die einzelnen Module – etwa für Datenerfassung, -vorverarbeitung, Modellgenerierung sowie die automatisierte V&V – werden zunächst unabhängig implementiert und anschließend in ein Gesamtsystem integriert.

Ein zentraler Bestandteil der methodischen Vorgehensweise ist die empirische Validierung des Frameworks durch eine praxisnahe Fallstudie. Hierbei werden reale Prozessdaten eines diskreten Materialflusssystems herangezogen, um die Simulationsergebnisse kontinuierlich mit den tatsächlichen Betriebsdaten abzugleichen. Die Evaluation erfolgt durch die Analyse quantitativer Metriken (z. B. Abweichungsmaße, Reaktionszeiten) und wird durch qualitative Analysen ergänzt, um systemische Fehlerquellen aufzudecken (**Uhlemann2017**).

Die Implementierung moderner Softwaretechnologien und datengetriebener Ansätze, wie etwa maschinelles Lernen, soll dabei helfen, die Validierungsprozesse zu automatisieren und flexibel an veränderte Betriebsbedingungen anzupassen. Der iterative Entwicklungsprozess ermöglicht es, das Framework kontinuierlich zu optimieren und auf die spezifischen Anforderungen unterschiedlicher industrieller Anwendungen zuzuschneiden. So wird ein dynamisches Instrument geschaffen, das nicht nur den theoretischen Anforderungen genügt, sondern auch in der Praxis anwendbar ist (**Tao2018; Kritzinger2018**).

Zusammenfassend bietet diese Arbeit einen innovativen Beitrag zur Weiterentwicklung digitaler Zwillinge in diskreten Materialflusssystemen. Durch die Kombination theoretischer Fundierung, moderner datengetriebener Methoden und empirischer Validierung wird ein Framework vorgestellt, das signifikante Verbesserungen in der Modellgenauigkeit und Effizienz von V&V-Prozessen ermöglicht. Die erzielten Ergebnisse tragen dazu bei, die Zuverlässigkeit digitaler Zwillinge zu erhöhen und bieten zugleich wertvolle Ansätze für die zukünftige Entwicklung in einer zunehmend digitalisierten und vernetzten Produktionslandschaft.





## Chapter 3

# Theoretical Foundations and State of the Art

### 3.1 Digital Twin: Definition and Concepts

#### 3.1.1 Types of Digital Twins

#### 3.1.2 Data-Driven Digital Twin (DT)

#### 3.1.3 Rationale from the Paper

### 3.2 Data-Driven Modeling: Principles and Methods

### 3.3 Material Flow Planning and Simulation

#### 3.3.1 Fundamental Concepts from Christian's Book

#### 3.3.2 Processes and Resources

#### 3.3.3 Insights from Pinedo's Book

#### 3.3.4 Production Planning and Control

### 3.4 Validation and Verification in Simulation

#### 3.4.1 Standard Formats and Data Sources for Simulation Models

#### 3.4.2 Validation Methods from Christian's Book

#### 3.4.3 Key Figures and KPIs

#### 3.4.4 Presentation of My Standard Format, *Pro Conta*, versus Other Formats

# **Chapter 4**

## **State of Research**

### **4.1 Existing Approaches for the Validation and Verification of Digital Twins**

### **4.2 Automatic Model Generation and Its Reasons/Challenges**

#### **4.2.1 Always Up-to-Date and Online Validation (Twin no longer represents reality)**

#### **4.2.2 High Complexity Requires Automated Procedures**

#### **4.2.3 Framework Promotes Standardization and Transparency**

#### **4.2.4 Scalability of Automated Approaches**

#### **4.2.5 Manual Validations May Contain Bias**

#### **4.2.6 Cost Savings**

### **4.3 Limitations of Current Standard Formats and Data Structures**

### **4.4 Gaps and Open Questions in Research**

# **Chapter 5**

## **Methodology and Framework Development**

### **5.1 Requirements Analysis**

### **5.2 Conceptualization**

#### **5.2.1 Description of the Framework for Automatic Model Generation**

#### **5.2.2 Verification of the Model: Structured Pipeline and Validation Goals**

#### **5.2.3 Data-Driven Validation: Event Data and KPIs for System Evaluation**

#### **5.2.4 Definition of Validation Scenarios (e.g., Masking, Blackbox Scenarios)**

#### **5.2.5 Automatic Evaluation Criteria for Model Validity (e.g., similarity to $R^2$ , Throughput Times, Efficiency Metrics)**

#### **5.2.6 Handling of Missing or Limited Data**

Waswani et al., 2017

# **Chapter 6**

## **Implementation of the Framework**

### **6.1 Technical Implementation of Model Generation**

#### **6.1.1 UML Diagram**

#### **6.1.2 Integration into Existing Systems**

### **6.2 Automatic Validation of the Generated Model**

### **6.3 Use of External Information Sources for Model Verification**

### **6.4 Interfaces to Material Flow Systems and Data Collection**

# **Chapter 7**

## **Case Study: Validation of a Digital Twin in the Production System**

### **7.1 Description of the Production System and Available Data**

#### **7.1.1 Description of the IoT Factory**

#### **7.1.2 Data Preparation and Analysis**

### **7.2 Construction of the Digital Twin for the Scenario**

### **7.3 Conducting the Validation Experiments**

### **7.4 Results and Interpretation: Limits and Errors of the Model**

### **7.5 Limits and Falsifiability of the Model Based on Real Data**

# **Chapter 8**

## **Discussion of the Results**

### **8.1 Critical Reflection on the Framework Development**

### **8.2 Limitations of Automated Validation**

### **8.3 Significance and Robustness of the Developed Methods**

#### **8.3.1 Comparison with Manual Modeling**

### **8.4 Implications for Research and Practice**

# **Chapter 9**

## **Conclusion and Outlook**

### **9.1 Summary of the Key Results**

### **9.2 Answering the Research Questions**

### **9.3 Outlook: Possible Further Development of the Framework**

### **9.4 Recommendations for Practical Application**



# Bibliography

- Abdoune, F., Nouiri, M., Cardin, O., & Castagna, P. (2023). Digital twin lifecycle: Core challenges and open issues. In T. Borangiu, D. Trentesaux, & P. Leitão (Eds.), *Service oriented, holonic and multi-agent manufacturing systems for industry of the future: Proceedings of sohoma 2022* (pp. 157–167). Springer International Publishing. [https://doi.org/10.1007/978-3-031-19647-3\\_14](https://doi.org/10.1007/978-3-031-19647-3_14)
- Arnold, D., & Furmans, K. (2005). *Materialfluss in logistiksystemen* (4th ed.). Springer-Verlag. <https://doi.org/10.1007/b139029>
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., et al. (2020). Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information fusion*, 58, 82–115.
- Frank, A. G., Dalenogare, L. S., & Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. *International Journal of Production Economics*, 210, 15–26. <https://doi.org/10.1016/j.ijpe.2019.01.004>
- Grieves, M. (2014). *Digital twin: Manufacturing excellence through virtual factory replication* (tech. rep.) (White Paper).
- Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterising the digital twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology*, 29, 36–52. <https://doi.org/10.1016/j.cirpj.2020.02.002>
- Judijanto, L., Qadriah, L., Prabowo, I. A., Widyatmoko, W., & Sabila, P. C. (2024). Trends in digital twin technology for industry 4.0: A bibliometric study. *The Eastasouth Journal of Information System and Computer Science*, 2(02), 92–104.
- Kovacs, G., & Kostal, P. (2016). Mathematical description of material flow. *Materials science and technology*, 1.

- Kritzinger, W., Karner, M., Traar, G., Henjes, J., & Sihn, W. (2018). Digital twin in manufacturing: A categorical literature review and classification. *IFAC-PapersOnLine*, 51(11), 1016–1022. <https://doi.org/10.1016/j.ifacol.2018.08.474>
- Lugaresi, G., & Matta, A. (2021). Automated digital twins generation for manufacturing systems: A case study. *IFAC-PapersOnLine*, 54(1), 749–754. <https://doi.org/10.1016/j.ifacol.2021.08.087>
- Oztemel, E., & Gursev, S. (2020). Literature review of industry 4.0 and related technologies. *Journal of intelligent manufacturing*, 31(1), 127–182.
- Robinson, S. (2014). *Simulation: The practice of model development and use*. Bloomsbury Publishing.
- Schwede, C., & Fischer, D. (2024). Learning simulation-based digital twins for discrete material flow systems: A review. *2024 Winter Simulation Conference (WSC)*, 3070–3081.
- Shao, G., Hightower, J., & Schindel, W. (2023). Credibility consideration for digital twins in manufacturing. *Manufacturing Letters*, 35, 24–28. <https://doi.org/10.1016/j.mfglet.2022.11.009>
- Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2018). Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology*, 94(9-12), 3563–3576. <https://doi.org/10.1007/s00170-017-0233-1>
- Thiede, S., Seow, Y., Andersson, J., & Johansson, B. (2013). Environmental aspects in manufacturing system modelling and simulation—state of the art and research perspectives. *CIRP Journal of Manufacturing Science and Technology*, 6(1), 78–87. <https://doi.org/10.1016/j.cirpj.2012.10.002>
- Trauer, J., Schweigert-Recksiek, S., Schenk, T., Baudisch, T., Mörtl, M., & Zimmermann, M. (2022). A digital twin trust framework for industrial application. *Proceedings of the Design Society*, 2, 293–302.
- Waswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *NIPS*.
- Wright, L., & Davidson, S. (2020). How to tell the difference between a model and a digital twin. *Advanced Modeling and Simulation in Engineering Sciences*, 7(1), 13. <https://doi.org/10.1186/s40323-020-00147-4>
- Zhang, L., Zhou, L., & Horn, B. K. (2021). Building a right digital twin with model engineering. *Journal of Manufacturing Systems*, 59, 151–164. <https://doi.org/10.1016/j.jmsy.2021.02.009>