Automatic Verification and Validation of Automatically Generated Simulation-Based Digital Twins for Discrete Material Flow Systems

---

Author: Daniel Fischer

Supervisor: Prof. Christian Schwede

Program: Research Master Data Science

Institution: Hochschule Bielefeld (HSBI)

Submission Date: March 16, 2025

Contents

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

ML Machine Learning. 1

SBDT Simulation-Based Digital Twin. 1

V&V Verification and Validation. 1

Chapter 1

Introduction

1.1 Initial Situation

Digital Twins (DT) are a key technology at the front of the fourth industrial revolution, coined Industry 4.0. The latter term is characterized by the interaction 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 this ideal by trying to remain competitive through the adoption of innovative technologies that promise enhanced productivity and reduced operational costs. One such technology that supports this transformation is the DT. It can be defined as a virtual representation of physical assets enabling real-time monitoring and optimization (Tao et al., 2018). The DT bridges 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, a unified definition of DT has yet to be established due to the diverse requirements and perspectives across different disciplines. In engineering, the focus might be on the real-time interaction between physical systems and their digital counterparts, whereas in computer science, the emphasis is often on data integration and simulation capabilities. These varying priorities result in multiple interpretations and applications of the term 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 terms used to describe 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).

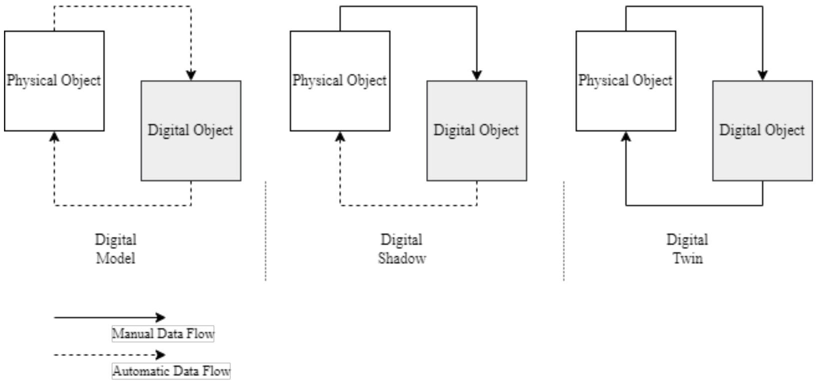

Figure 1.1: Comparison of Digital Shadow (DS), Digital Model (DM) and Digital Twin (DT) as presented by Kritzinger (2018). This distinction is crucial for understanding validation requirements across different digital representation types.

The DM represents the most basic form. It contains manual data connections between 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 frequently used for simulation purposes and is often misclassified as a DT in the literature. 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. Manufacturing stands out due to its high potential for process optimization and automation. This thesis focuses on the latter, 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, in accordance to 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 representations 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). A second driver of DT innovation lies in the widely available data from IoT devices and sensors. Unlocking these potentials in model training and real-time adaption of the DT is crucial for its modelling capabilities (Tao et al., 2018).

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 (Schwede & Fischer, 2024). 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 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 throughout the whole system lifecycle (Boschert & Rosen, 2016; Lim et al., 2020).

1.2 Problem

Despite the transformative potential of DT, their implementation can be challenging. The creation and maintenance of accurate Digital Twins require substantial investments 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. Automatic generation may be an elegant solution, but possesses the risk of overfitting or biased predictions (Geman et al., 1992). Manufacturing data used as training data must be preprocessed and cleaned thoroughly. The DT per definition has to be able to perform real-time decision making without excessive time lags. This yields the requirement to extract, transform, clean and load (ETL process, Vassiliadis et al., 2002) data on the fly, performing inference fully in memory. All these hurdles are challenges to automatic learning (Ribeiro et al., 2016; Zhao et al., 2024). As industries integrate DT into their production processes, establishing trust becomes fundamental as well (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 (Lapointe & Rivard, 2005).

Even if DT learning is successfully performed, questions regarding its correctness, precision, and robustness remain unanswered. These questions are tackled by validation, verification and uncertainty quantification frameworks (VVUQ) (Sel et al., 2025). 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 (Bitencourt et al., 2023; Hua et al., 2022). This creates inefficiencies, particularly in the context of automated

DT generation, where such manual processes conflict with the goal of reducing development effort. Hua et al., 2022 even argue that there are no robust and standardized VVUQ methods for Digital Twins. One hurdle to standardized VVUQ frameworks is the lack of a clear definition for validity and verification in the context of Digital Twins (Bitencourt et al., 2023).

For discrete material flow systems, these challenges are even more pressing due to their processual nature and inherent stochastic elements. Manufacturing processes may fail due to anomalies, resource constraints or human error. VVUQ approaches have to anticipate these risks. When DTs for these systems are generated automatically, conventional validation approaches become problematic, as they negate much of the efficiency gained through automation. This creates a fundamental contradiction: While automated Digital Twin generation reduces initial development- and updating efforts, it simultaneously increases the complexity of validation and verification, potentially counteracting its intended efficiency gains.

1.3 Objective

The thesis thus addresses this contradiction by developing a data-driven framework for automated VVUQ of automatically generated, simulation-based DTs which have been learned from data. The focus lies on DMFS due to their practical relevance and dynamical, processual nature. The endeavor can further be concretized by the following research questions (RQ):

- **RQ1:** How can automated validation and verification processes for Digital Twins be efficiently implemented to maintain both accuracy and computational feasibility?
- **RQ2:** Which data-driven approaches are best suited to identify discrepancies between simulated behavior and real operational data in discrete material flow systems?
- **RQ3:** To what extent does the developed framework improve the quality and reliability of digital twins compared to traditional VV methods?

This thesis addresses these questions by proposing that object-centric event logs—the same data structures often used to generate DT in manufacturing—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. The development and monitoring of generic, statistically grounded reference values is a key aspect of this approach. Such key indicators need to have an underlying distribution and have to be quantifiable. The framework will be evaluated using a case study from the discrete material flow domain, providing empirical evidence of its effectiveness in improving model accuracy and efficiency.

1.4 Structure and Methodology

1.4.1 Structure

The thesis is structured into eight chapters. Initially, theoretical foundations and definitions (e.g., DT, VVUQ, process mining, DMFS) are established. After that, the literature on VVUQ and related methodologies is critically reviewed. The methodological approach, based on design science research (DSR), is then presented, followed by the detailed implementation of the framework and its evaluation via a case study. Finally, the discussion and conclusions summarize the research findings and outline future research directions.

Chapter 2 establishes the theoretical groundwork by delivering a clear understanding and definitions of the topics. These definitions include the topics DT, VVUQ, process mining and DMFS. The chapter follows an hourglass scheme: The first topic includes an overview of different DT and their applications is given, narrowing down on data-driven DT and their daughters, automatically generated DT which stem from gathered data. The first topic closes with a differentiation of SBDT and classical simulation approaches. The second topic focusses on an verification and validation in general, not drawing an immediate connection between them. Historical development of verification and validation are shown. An examplary continuous process is described. Furthermore, VVUQ in the context of AMG and SBDT is presented. The third topic introduces process mining (PM) and its data formats and tools. A connection between PM and VVUQ is drawn. The chapter concludes with a summary of material flow planning and simulation, which is the domain of the case study.

Chapter 3 educates the reader by explaining the landscape of VVUQ in literature. Relevant VVUQ approaches are described and critically evaluated. The chapter serves as a short literature review by further shedding light on AMG for DT. Relevant problem fields in practice are examined, further specifying the problem niche of the thesis. ML-based approaches for general VVUQ of models, not necessarily DT, are presented. The chapter further clarifies the research gap and reviews the current state of research in the field of DTs, focusing on validation and verification. The chapter provides an overview of the existing literature on the topic, highlighting the gaps and challenges in the current research.

Chapter 4 outlines the methodology used to develop the proposed framework. It begins with a classical requirements analysis, deriving functional, technical, and data format requirements from theoretical findings. The chapter elaborates on the data-based validation strategy, machine learning-based validation approach, metrics for model evaluation, and online validation with continuous monitoring.

Chapter 5 presents the framework implementation, starting with the architecture and system setup, followed by detailed descriptions of event log processing, simulation integration, and the machine learning pipeline.

Chapter 6 presents the case study results, evaluating the framework's effectiveness in improving Digital Twin quality. It describes the application scenario and data basis, the automatically generated DT, validation experiments, and result interpretation. It concludes with a comparison to manual validation methods.

Chapter 7 discusses the implications of the results and provides recommendations for future research. It evaluates the framework in light of the research questions, examines the significance of verification in automatically generated DT, addresses limitations, and explores implications for research and practice.

Chapter 8 concludes the thesis by summarizing key findings and their implications. It addresses the research questions and hypotheses, discusses the results' significance, acknowledges limitations, and provides recommendations for future work.

1.4.2 Methodology

Methodologically the thesis follows a design science research approach (DSR). This approach is characterized by the development of artifacts to solve practical problems (Hevner et al., 2004; Peffers et al., 2007). Artifacts in the sense of DSR are created objects or constructs which adress the given problem and contribute in direction of both theory and practice. The artifacts are evaluated in a real-world context to demonstrate their effectiveness. The thesis follows DSR by applying the cyclical DSR model, see figure 1.2.

The research paradigm of the thesis can be characterized as deductive-theory critical (Eberhard, 1987). A conceptual VVUQ framework is developed based on existing theoretical foundations, while deriving new requirements thorugh a re-

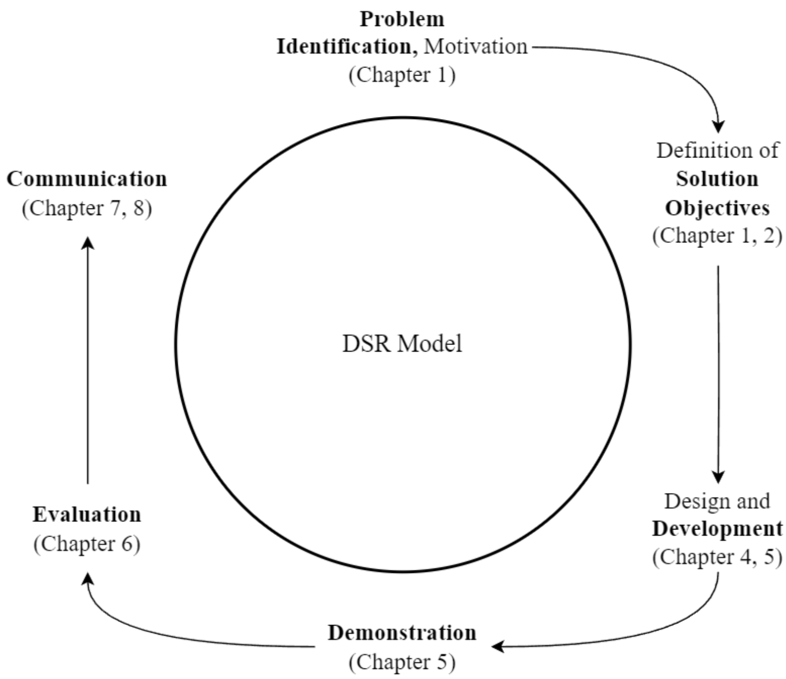

Figure 1.2: The cyclical design science research model. The model consists of six steps. The problem identification (1) refers to the research gap in automated VVUQ of SBDT. Defining the solution objectives (2) specifies the research gap by formulating questions and hypotheses based on the theoretical foundations. The design and development (3) phase includes the development of the framework. The demonstration (4) phase shows the application of the framework in a case study. The evaluation (5) phase assesses the effectiveness of the framework. The communication (6) phase concludes the research by presenting the results.

quirements analysis. The framework is then applied in a case study to evaluate its effectiveness. The research is critical in the sense that it aims to improve the efficiency and effectiveness of VVUQ for automatically generated DTs. Elements of empirical research are included through the case study and the data-driven approach.

Chapter 2

Theoretical Foundations and Status Quo

2.1 Digital Twin: Definition and Concepts

2.1.1 Types of Digital Twins

2.1.2 Data-Driven Digital Twins

2.1.3 Automatically Generated Digital Twins

2.1.4 Definitions and Differences from Classical Simulation Literature

2.2 Verification and Validation in the Context of Simulation-Based Digital Twins

2.2.1 Historical Development of V&V Concepts

2.2.2 Peculiarities of Automatically Generated Models

2.2.3 Theoretical Argumentation for Merging Verification and Validation

2.2.4 V&V as a Continuous Process

2.3 Process Mining and Event Logs

2.3.1 Object-Centric Event Logs as a Data Basis

2.3.2 Standard Formats and Their Importance for Automated Validation

2.3.3 Process Mining as a Bridge between Process Data and Model Validation

2.4 Material Flow Planning and Simulation

Chapter 3

State of Research

3.1 Existing Approaches to Validation and Verification of Digital Twins

3.2 Automatic Model Generation for Digital Twins

3.3 Machine Learning-Based Approaches for Model Validation

3.4 Research Gaps and Open Questions

Chapter 4

Conception of a Framework for Automated Validation

4.1 Requirements Analysis

4.2 Data-Based Validation Strategy

4.3 Machine Learning-Based Validation Approach

4.4 Metrics and Key Figures for Model Evaluation

4.5 Online Validation and Continuous Monitoring

Chapter 5

Implementation of the Framework

Chapter 6

Case Study: Validation in Practice

6.1 Application Scenario and Data Basis

6.2 Automatically Generated Digital Twin

6.3 Validation Experiments

6.4 Results and Interpretation

6.5 Comparison with Manual Validation Methods

Chapter 7

Discussion of the Results

7.1 Evaluation of the Developed Framework

7.2 Significance of Verification in Automatically Generated Digital Twins

7.3 Limitations of Automated Validation

7.4 Implications for Research and Practice

Chapter 8

Conclusion and Outlook

8.1 Summary of the Key Findings

8.2 Methodological and Theoretical Insights

8.3 Outlook

8.4 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

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.

Bitencourt, L., Silva, M., & Costa, J. (2023). Building trust in digital twin through verification and validation. *Proceedings of the 2023 International Conference on Digital Twin Technologies*, 55–63. https://doi.org/10.1000/bitencourt.2023.001

Boschert, S., & Rosen, R. (2016). Digital twin—the simulation aspect. *Mechatronic futures: Challenges and solutions for mechatronic systems and their designers*, 59–74.

Eberhard, K. (1987). *Einführung in die erkenntnis-und wissenschaftstheorie: Geschichte und praxis der konkurrierenden erkenntniswege*. Kohlhammer.

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

Geman, S., Bienenstock, E., & Doursat, R. (1992). Neural networks and the bias/variance dilemma. *Neural Computation*, *4*(1), 1–58. https://doi.org/10.1162/neco.1992.4.1.1

Grieves, M. (2014). *Digital twin: Manufacturing excellence through virtual factory replication* (tech. rep.) (White Paper).

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 75–105.

Hua, E. Y., Lazarova-Molnar, S., & Francis, D. P. (2022). Validation of digital twins: Challenges and opportunities. *2022 Winter Simulation Conference (WSC)*, 2900–2911. https://doi.org/10.1109/WSC57314.2022.10015420

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

Lapointe, L., & Rivard, S. (2005). A multilevel model of resistance to information technology implementation. *MIS quarterly*, 461–491.

Lim, K. Y. H., Zheng, P., & Chen, C.-H. (2020). A state-of-the-art survey of digital twin: Techniques, engineering product lifecycle management and business innovation perspectives. *Journal of Intelligent Manufacturing*, *31*(6), 1313–1337.

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.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, *24*(3), 45–77.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). " why should i trust you?" explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 1135–1144.

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.

Sel, K., Hawkins-Daarud, A., Chaudhuri, A., Osman, D., Bahai, A., Paydarfar, D., Willcox, K., Chung, C., & Jafari, R. (2025). Survey and perspective on verification, validation, and uncertainty quantification of digital twins for precision medicine. *npj Digital Medicine*, *8*(1), 40.

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.

Vassiliadis, P., Simitsis, A., & Skiadopoulos, S. (2002). Conceptual modeling for etl processes. *Proceedings of the 5th ACM international workshop on Data Warehousing and OLAP*, 14–21.

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

Zhao, Y. F., Xie, J., & Sun, L. (2024). On the data quality and imbalance in machine learning-based design and manufacturing—a systematic review. *Engineering*.