# A Vision for Advancing Digital Twins Intelligence: Key Insights and Lessons from Decades of Research and Experience with Simulation

Sanja Lazarova-Molnar[1,2,a]
[1]Institute of Applied Informatics and Formal Description Methods, Karlsruhe Institute of Technology, Karlsruhe, Germany
[2]Mærsk Mc-Kinney Møller Institute, University of Southern Denmark, Odense, Denmark

Keywords: Digital Twins, Simulation, Key Considerations, Goal-Oriented, Fusion of Data and Expert Knowledge.

Abstract: Digital Twins have revolutionized the domain of Modeling and Simulation by making use of the growing and cost-efficient possibilities to extract data from systems, as well as the increasing computational power. At the same time, Digital Twins have enabled tremendous advances in diverse cyber-physical systems by enabling better monitoring, predictive maintenance, design optimization, and informed decision-making. As their popularity evolved, the understanding of what a Digital Twins has become more and more dispersed and unclear. Here, we offer understanding of what a Digital Twin is based on our experience in research within its native domain of Modeling and Simulation, with a concrete focus on the key considerations that need to be made when developing Digital Twins or working with them. We, furthermore, emphasize the need to include all available knowledge for better-informed Digital Twins. To illustrate our ideas and vision, we use case studies from our research.

## 1 INTRODUCTION

Digital Twins have emerged as indispensable assets for industries seeking to optimize operations and enhance decision-making processes. At their core, Digital Twins are dynamic digital counterparts of physical entities, continuously updated based on real-world data, simulating and mirroring physical systems' behavior (Friederich et al., 2022; Grieves, 2014). Digital Twins are typically realized through computational models like machine learning or simulation models (Masood & Sonntag, 2020). These twins, both digital and physical, are interconnected in near-real-time, facilitating constant information exchange and synchronization. They are inherently data-driven, relying heavily on supervised and unsupervised machine learning techniques (Friederich et al., 2022).

In this paper, we offer understanding of what a Digital Twin is based on our decades-long experience in research within its native domain of Modeling and Simulation (Friederich & Lazarova-Molnar, 2023; Lazarova-Molnar & Horton, 2003, 2004), with a concrete focus on the key considerations that need to be made when developing Digital Twins or working with them. Drawing from this experience, we emphasize on the Digital Twins' goal-oriented nature, vital to comprehend for laying the groundwork for exploring their Digital Twins' applications and operationalization. Thus, we concentrate and elaborate on the imperative of goal-oriented digital twin design. Besides other key lessons drawn from the domain of Modeling and Simulation, we also present our vision of how to enable better-informed Digital Twins by enabling systematic inclusion of all information that we have available. To this point, we briefly introduce two case studies that illustrate this vision and draft our future research.

## 2 KEY CONSIDERATIONS FOR DIGITAL TWINS

Drawing from our experience in developing methods for Digital Twins and our work in Modeling and Simulation (Francis et al., 2021; Hua et al., 2022; Mohamed et al., 2023), especially in the context of data-driven simulation (Lazarova-Molnar & Li, 2019), we have developed a number of key considerations that we elaborate in the following:

[a] https://orcid.org/0000-0002-6052-0863



Lazarova-Molnar, S.
A Vision for Advancing Digital Twins Intelligence: Key Insights and Lessons from Decades of Research and Experience with Simulation.
DOI: 10.5220/0012884800003758

In *Proceedings of the 14th International Conference on Simulation and Modeling Methodologies, Technologies and Applications (SIMULTECH 2024)*, pages 5-10
ISBN: 978-989-758-708-5; ISSN: 2184-2841

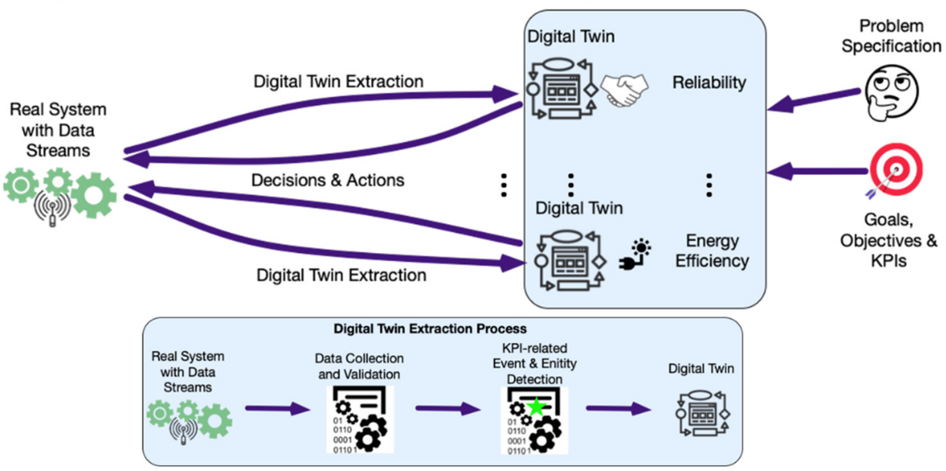


Figure 1: Goal-oriented Digital Twins.

- Goal-oriented Digital Twins: The first two steps in devising classic simulation studies are *Problem Specification* and *Objective Specification* (Lazarova-Molnar & Li, 2019). Thus, when building simulation models, one has to know what problem the models should solve and set the objectives to target that. This is pretty much the same with Digital Twins, which also need to be built to solve a specific problem or a set of problems. There cannot be a Digital Twin that fully models all aspects of a given system and answers all possible questions. As such, we have to talk about goal-oriented Digital Twins. In this manner, a system does not have a single Digital Twin, but a set of Digital Twins that are projections of a system's behavior on predefined problems and objectives. This point is illustrated in Figure 1, where we see that one Digital Twin is needed to target energy efficiency of a system, and another one for reliability analysis, for instance. These Digital Twins also need different data streams and have different underlying models. In additon, the underlying modeling formalisms might be different as well. They need to be, however, driven by the problem specifications.
- Digital Twins necessitate a feedback loop: Digital Twins are different from simulation models in that that their underlying models update in near-real-time, continuously reflecting changes that occur in the corresponding real systems. In addtion, Digital Twins need to deliver simulation-informed decisions back to the real system. For this, Digital Twins need to incorporate a feedback loop that enables the bidirectional communication process which is the detail that makes Digital Twins different from traditional and static simulation models.
- Digital Twins automate simulation modeling: Digital Twins have emerged as a result of the large prevalence of data and computing resources. This provided a chance to automate some of the most expensive computational techniques, i.e., in this case modeling and simulation. Thus, Digital Twins can be seen as an automation of the classical modeling and simulation processes, and as such, need to utilize significant portion of the formal background that has been developed in this area. This formal background has been around for many decades and it has been mainly targeted at knowledge-driven model development (Law et al., 2007; Maria, 1997; Zeigler et al., 2000). This body of knowledge can only benefit, clarify and support the understanding and advancement of Digital Twins.
- Automatically deriving simulation models from ongoing data collection, as is the case with Digital Twins, provides opportunity to simultaneously perform validation, and, thus, release the need for explicit and separate validation processes. Moreover, validation of Digital Twins' underlying simulation models is no longer additional optional step. On contrary, integrated validation becomes a prerequisite and an integral part of Digital Twins as ensuring that the underlying model accurately reflects the real system is necessary for a model to be quialified as a Digital Twin model

(Friederich et al., 2022; Hua et al., 2022; Zare & Lazarova-Molnar, 2024).

- Integration of all that we know or will know: Digital Twins need not throw away what we already know or may know in future. Thus, Digital Twins need to be built in a way that they seamlessly integrate all that we know or might know in future about the systems of interest. This knowledge can come in many different forms, i.e., as physics-based models, or as expert knowledge from experience. This is a strong point in our line of research, on which we elaborate more in the following.

To illustrate our understanding and definition of Digital Twins, we designed and further developed a framework that we initially presented in (Friederich et al., 2022). For this contribution, we further iterated on the framework, to present an updated version that is more complete and considers the elements that are used the derive decisions, e.g., Data Analytics, Optimization, etc. This updated framework is illustrated in Figure 2.

## 3 FUSION OF DATA AND EXPERT KNOWLEDGE FOR DIGITAL TWINS

Development of Digital Twins to accurately represent their real-world physical counterparts poses significant challenges. In academic literature, two main approaches to modeling Digital Twins emerge, each adhering to distinct paradigms. The first approach relies on traditional knowledge-driven modeling methods, while the second favors newer data-driven model extraction strategies. Presently, there is a notable preference for data-driven techniques, which reduces reliance on human expertise. However, solely relying on a data-driven approach neglects the potential benefits of integrating expert knowledge.

In the last decade, along the same line of thought, we have been fully immersed in the data-driven methods and discoveries, slightly ignoring the fact that we, humans, already do know a lot about these systems that we model and simulate. For this reason, it is essential that we develop methodology that can seamlessly integrate human/expert knowledge in the models that we extract. An engineer that has been working with a given system, knows a lot about this system. We should find a way to utilize that and not simply override with data. Furthermore, there have been centuries of development of physics-based models for various phenomena. Therefore, it is imperative to develop methods to interface this knowledge with data, as well as use both information sources to complement and cross-validate each other.

Expert knowledge complements and enhances data-driven methodologies by offering insights that may be challenging to obtain through data alone, or at least to obtain within a given time range. Additionally, expert knowledge provides nuanced understandings of phenomena based on expert experiences and contexts, thereby addressing other challenges associated with data-driven Digital Twins model extractions, such as data scarcity. As a result,

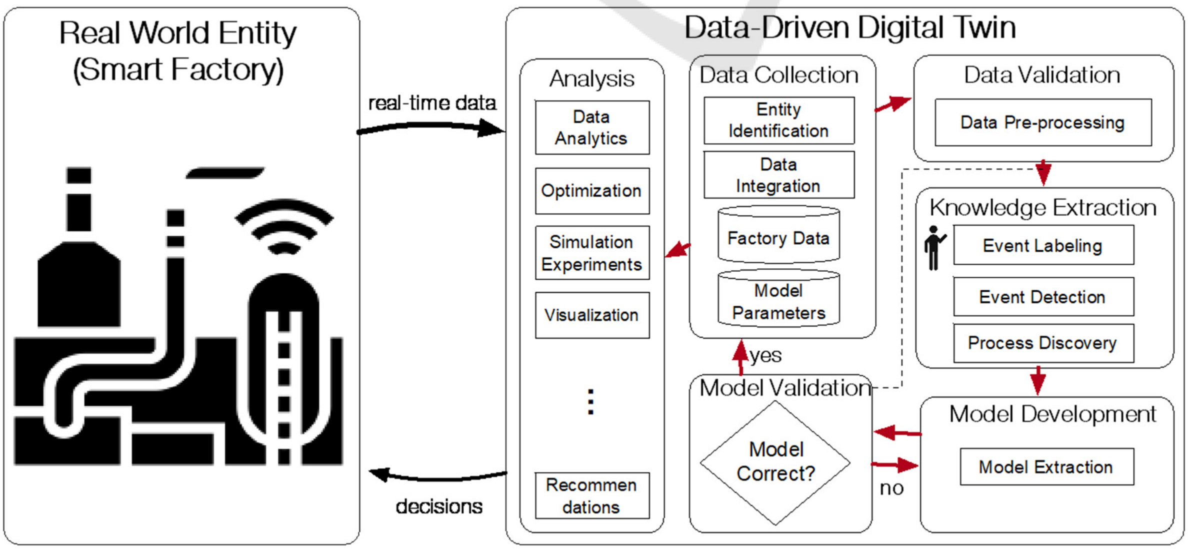


Figure 2: Framework for Data-Driven Digital Twins, extension from (Friederich et al., 2022).

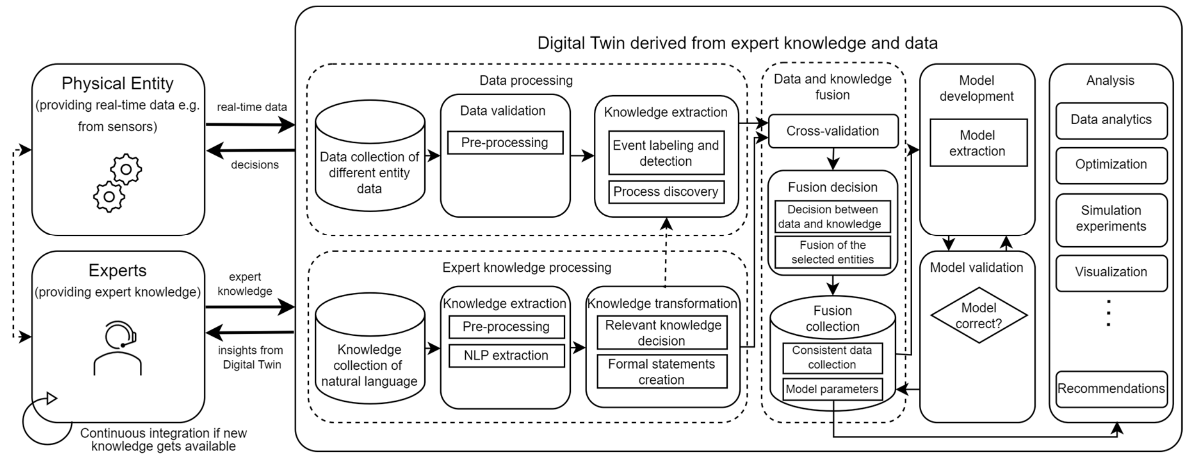


Figure 3: Fusion DT Framework derived from both expert knowledge and data, introduced in (Michelle Jungmann & Sanja Lazarova-Molnar, 2024).

combining data and expert knowledge can both enable more accurate and more efficient Digital Twin model development.

To enhance the aforementioned Digital Twins' full potential, a framework that seamlessly integrates human intelligence with technological innovation is essential. Thus, we have also recently presented an initial framework for Digital Twins (Michelle Jungmann & Sanja Lazarova-Molnar, 2024), which we illustrate in Figure 3. We intend to further enhance this framework and methodology to leverage advanced technologies such as large language models to streamline digital twin workflows and augment decision-making capabilities. By emphasizing the importance of visualization in facilitating knowledge integration, we outline a roadmap for advancing digital twin performance.

To illustrate the framework for integration of expert knowledge in Digital Twins, we developed case studies that we briefly illustrate in the following section, and refer to the corresponding sources for more detailed decsriptions.

In the following, we briefly refer to two of our latest case studies to illustrate the idea of including expert knowledge for automated extraction of simulation models, as a vision toward Digital Twins that utilize the same information sources. Both studies focus on reliability modeling and analysis.

### 3.1 Extraction of Fault Tree Models from Fusion of Data and Expert Knowledge

Grounded in our empirical evidence, this case study offers a demonstration of fusing digital twin efficacy through data-driven methodologies and human expertise. We designed the process of fault tree extraction from a fusion of time series data streams and expert knowledge. With this case study, we demonstrated the potential of using expert knowledge for better-informed digital twins in enhancing system reliability and safety (Niloofar & Lazarova-Molnar, 2021).

The workflow, that is fully detailed in (Niloofar & Lazarova-Molnar, 2021), consists of using both data and expert knowledge statements to extract fault tree models that can be utilized for assessing reliability of cyber-physical systems. With this, we were able to enhance the synergy between human expertise and sensor data by refining methods to translate expert insights into data-mergeable formats. We, furthermore, introduced Bayesian probability updating, weighted cut sets, and validated expert knowledge through incremental data gathering. In the upcoming work, we focus on generalizing systematic integration of data-driven and knowledge-driven fault tree analysis.

### 3.2 Extraction of Petri Net Models from Data and Expert Knowledge

In a more recent effort that was initiated by the work in (Michelle Jungmann & Sanja Lazarova-Molnar, 2024) and further expanded in (Michelle Jungmann & Sanja Lazarova-Molnar, 2024), we extracted reliability-centred Petri nets by fusing streaming data from a system and expert knowledge. The results we obtained were promising, demonstrating the feasibility of explicitly combining expert knowledge and data for more accurate reliability models. With this, we highlighted the value that expert knowledge can bring in faster extraction of more accurate models.

However, we also encountered challenges in these fusion processes, where one of the major ones was – who do we trust more when data and expert knowledge contrast each other. For fusing data and expert knowledge, we introduced two main strategies, termed *a priori* and *a posteriori*. We implemented these fusion algorithms in a case study on reliability modeling, where we formulated and formalized four expert knowledge statements that varied in complexity. Synthetic data was generated for the *a priori* fusion algorithms, and the results of each fusion algorithm were executed and analyzed, leading to the extraction of a Petri net model. The case study demonstrated the potential of combining expert knowledge and data for different types of reliability information using the proposed strategies. In the upcoming work, we will involve refining the approach by improving the quality and integration of expert knowledge, automating formalization from natural language, incorporating fuzzy logic, and addressing data gaps in logs.

# 4 CONCLUSION

Addressing the evolving landscape of Digital Twins requires a proactive approach to overcoming challenges and charting future directions. Drawing from our long-time experience of research in the domain of Modeling and Simulation, we identified and elaborated on key considerations for developing Digital Twins. We, furthermore, emphasized the importance of developing goal-oriented Digital Twins, as well as utilizing all available knowledge in a systematic way to enhance more accurate underlying models of Digital Twins, and with this enhance their intelligence. We illustrated our key points by case studies from our research, which we used to show directions for our future research.

In conclusion, enhancing Digital Twins' usefulness and performance requires a holistic approach that integrates human intelligence, data analytics, and practical problem-oriented approach, which also builds upon the existing body of knowledge from the well-established domain of Modeling and Simulation. With this, we can enhance Digital Twins efficacy and their utilization across diverse industries and for different goals.

# ACKNOWLEDGEMENTS

The authors extend their thanks for the funding received from the ONE4ALL and DMaaST projects funded by the European Commission, Horizon Europe Programme under the Grant Agreements No. 101091877 and No. 101138648, correspondingly.

# REFERENCES

Francis, D. P., Lazarova-Molnar, S., & Mohamed, N. (2021). Towards data-driven digital twins for smart manufacturing. Proceedings of the 27th International Conference on Systems Engineering, ICSEng 2020,

Friederich, J., Francis, D. P., Lazarova-Molnar, S., & Mohamed, N. (2022). A framework for data-driven digital twins of smart manufacturing systems. *Computers in Industry*, *136*, 103586.

Friederich, J., & Lazarova-Molnar, S. (2023). A Framework for Validating Data-Driven Discrete-Event Simulation Models of Cyber-Physical Production Systems. 2023 Winter Simulation Conference (WSC),

Grieves, M. (2014). Digital twin: manufacturing excellence through virtual factory replication. *White paper*, *1*(2014), 1-7.

Hua, E. Y., Lazarova-Molnar, S., & Francis, D. P. (2022). Validation of digital twins: challenges and opportunities. 2022 Winter Simulation Conference (WSC),

Jungmann, M., & Lazarova-Molnar, S. (2024). Fusing Expert Knowledge and Data for Simulation Model Discovery in Digital Twins: A Case Study from Reliability Modeling. Winter Simulation Conference 2024,

Jungmann, M., & Lazarova-Molnar, S. (2024). Towards Fusing Data and Expert Knowledge for Better-Informed Digital Twins: An Initial Framework. *Procedia Computer Science*.

Law, A. M., Kelton, W. D., & Kelton, W. D. (2007). *Simulation modeling and analysis* (Vol. 3). Mcgraw-hill New York.

Lazarova-Molnar, S., & Horton, G. (2003). An Experimental Study of the Behaviour of the Proxel-Based Simulation Algorithm. SimVis,

Lazarova-Molnar, S., & Horton, G. (2004). Proxel-based simulation of a warranty model. European Simulation multiconference,

Lazarova-Molnar, S., & Li, X. (2019). Deriving simulation models from data: steps of simulation studies revisited. 2019 Winter Simulation Conference (WSC),

Maria, A. (1997). Introduction to modeling and simulation. Proceedings of the 29th conference on Winter simulation,

Masood, T., & Sonntag, P. (2020). Industry 4.0: Adoption challenges and benefits for SMEs. *Computers in Industry*, *121*, 103261.

Mohamed, N., Lazarova-Molnar, S., & Al-Jaroodi, J. (2023). Digital Twins for Energy-Efficient Manufacturing. 2023 IEEE International Systems Conference (SysCon),

Niloofar, P., & Lazarova-Molnar, S. (2021). Fusion of data and expert knowledge for fault tree reliability analysis of cyber-physical systems. 2021 5th International Conference on System Reliability and Safety (ICSRS),

Zare, A., & Lazarova-Molnar, S. (2024). Validation of Digital Twins in Labor-Intensive Manufacturing: Significance and Challenges. The 7th International Conference on Emerging Data and Industry 4.0,

Zeigler, B. P., Praehofer, H., & Kim, T. G. (2000). *Theory of modeling and simulation*. Academic press.