Model-Driven Software Engineering to Foster the Adoption of Digital Twins

Arturo Barriga[1]

[1]Quercus Software Engineering Group (https://quercusseg.unex.es), Departamento de Ingeniería Sistemas Informáticos y Telemáticos, Universidad de Extremadura, Avenida de la Universidad s/n, Cáceres, 10003, Spain

arturobc@unex.es

Abstract. *Digital twins are virtual replicas of physical systems that offer significant benefits in terms of productivity and sustainability. However, most digital twin solutions are tailored to specific use cases, and there is a lack of domain-independent development frameworks. As a result, developers must build each system from scratch, which increases project timelines and costs and hinders the scalability of solutions. This ongoing doctoral project proposes a Model-Driven Software Engineering (MDSE) approach to simplify the development of digital twins. Ultimately, the presented approach aims to promote the widespread adoption of digital twins across sectors, making their development more accessible, scalable, and cost-effective.*

1. Introduction

A digital twin is a virtual representation of a physical system that is continuously updated and synchronized with its real counterpart via real-time data streams [Minerva et al. 2020, Sharma et al. 2022]. By integrating technologies such as the Internet of Things (IoT) and machine learning, digital twins create a dynamic, data-driven environment that allows users to simulate scenarios, optimize performance, and proactively identify potential failures or undesirable outcomes before they occur [Rathore et al. 2021, Segovia and Garcia-Alfaro 2022, Minerva et al. 2020]. Moreover, digital twin applications span numerous industries, including manufacturing [Friederich et al. 2022], smart cities [Wang et al. 2023], and energy [Stadtmann et al. 2023]. In this regard, digital twin systems offer substantial benefits, such as improving efficiency and sustainability, and supporting informed decision-making.

However, a major challenge in digital twin development is the lack of domain-independent frameworks [Macías et al. 2023, Mihai et al. 2022, Fuller et al. 2020], which forces developers to build custom solutions for each specific use case. This lack of standardization means that each digital twin must be developed from scratch, which leads to a significant increase in both time and cost and limits the scalability of systems [Lehner et al. 2021, Vale et al. 2023]. This not only raises the barriers to entry for organizations but also hinders the widespread adoption of digital twin technology.

To address these issues, this ongoing doctoral project proposes a Model-Driven Software Engineering (MDSE) approach that includes an abstract syntax, a concrete graphical syntax, and Model-to-Text (M2T) transformations to facilitate the development of digital twins. In this sense, MDSE is a software development methodology that focuses on the use of models as central artifacts throughout the development process [Brambilla et al. 2017]. By using high-level models to represent system structure and behavior, along with M2T transformations, MDSE enables the automatic generation of executable code and other necessary artifacts directly from the models [Hailpern and Tarr 2006, Stahl et al. 2006]. This approach reduces the need for manual coding, minimizes human errors resulting from laborious tasks, and speeds up the development process [Brambilla et al. 2017, Hailpern and Tarr 2006, Stahl et al. 2006]. As a result, developers can focus on refining the logic and functionality of the system instead of dealing with repetitive manual tasks, which speeds implementation and improves consistency throughout the system.

In this context, the proposed MDSE approach involves defining the conceptual model of digital twins from a domain-independent perspective, identifying the essential properties, components, and functionalities of various digital twins documented in the literature. This conceptual model will then be formalized through a metamodel, specifying the rules and semantics governing the structure and behavior of valid digital twin systems. To facilitate intuitive model creation, a graphical editor will be implemented, allowing users to design digital twin systems without requiring extensive programming knowledge. Finally, M2T transformations will be developed to automate the generation of executable code from the models, thereby significantly reducing manual coding efforts [Hailpern and Tarr 2006, Stahl et al. 2006].

Therefore, this project aims to encourage the widespread adoption of digital twins by making their development process more consistent, accessible, scalable, and cost-effective. By simplifying the design, validation, generation, and deployment phases, this approach not only promises to reduce development time and costs but also to promote wider use of digital twin technology across diverse industries, ultimately driving innovation and improving operational efficiency.

It should be mentioned that, although the proposed MDSE approach is domain-independent and could be applied to a variety of industries, this doctoral project pays special attention to the agricultural sector due to its important role in Extremadura, the author's region. In this regard, to validate the applicability of the proposed solution, this doctoral project also aims to implement agricultural digital twins in Extremadura, thus validating the proposed approach in real-world use cases.

The remainder of this manuscript is organized as follows. Section 2 presents an explicit formulation of the research questions and objectives of the doctoral project. Section 3 provides a summary of the current state of knowledge in the problem domain. Section 4 outlines the research methodology employed throughout the project. Section 5 presents the proposed solution and discusses the results obtained so far. Finally, Section 6 presents some of the challenges faced and outlines the next steps planned for the project.

2. Research Questions and Objectives

Given the context and background that justifies the development of this proposal, the main Research Questions (RQs) guiding the doctoral project are:

- **RQ1:** How can a conceptual model of digital twins be defined to effectively capture their essential components, features, and functionalities in a domain-independent manner?
- **RQ2:** How can MDSE be applied to the definition, validation, generation, and deployment of digital twins across industries?

- **RQ3:** What are the challenges of applying MDSE to digital twin development, and how can they be addressed?
- **RQ4:** How can the proposed MDSE approach be validated to assess its applicability and effectiveness in real-world digital twin implementations?

These research questions have been carefully selected. RQ1 is formulated to establish a foundational understanding of digital twins, ensuring that the proposed approach is applicable across multiple domains. RQ2 focuses on how MDSE methodologies can be applied to digital twin development, addressing the limitations of current ad-hoc implementations. RQ3 examines the challenges encountered when applying MDSE to digital twin development, identifying potential obstacles and proposing solutions to ensure the effectiveness of the methodology. Finally, RQ4 ensures that the proposed methodology is validated in practical use cases.

To address these research questions, the doctoral project aims to achieve the following objectives:

- **Ojective 1:** Define a domain-independent conceptual model that identifies the key characteristics, components, relationships, and functionalities of digital twins.
- **Ojective 2:** Propose and implement an MDSE approach to streamline the definition, validation, implementation, and deployment of digital twins across diverse industries.
- **Ojective 3:** Validate the applicability of the proposed model-driven approach in real-world case studies, with special attention to the agricultural sector.
- **Ojective 4:** Disseminate the findings and results of this research through presentations at relevant conferences and publications in high-impact journals.

The following section describes existing MDSE approaches to digital twin development, identifying their limitations and areas for improvement.

3. Current Knowledge and State of Existing Solutions

Traditionally, MDSE has been successfully applied to fields closely related to digital twins, such as the Internet of Things (IoT) and Cyber-Physical Systems (CPS) [Ciccozzi and Spalazzese 2016, Moin et al. 2020, Barriga et al. 2021, Barriga et al. 2023, Ihirwe et al. 2020]. Then, with the rise of digital twin technology, researchers have begun investigating the application of MDSE in this domain [Bibow et al. 2020, Dalibor et al. 2020, Vale et al. 2023], aiming to streamline the development process, reduce its complexity, and enhance automation.

In this regard, [Bibow et al. 2020] proposes an MDSE approach for creating reactive digital twins by defining event-driven rules. This work was later extended by [Dalibor et al. 2020] to include a graphical control interface (cockpit) for monitoring and interacting with physical entities. However, these solutions primarily focus on predefined rule-based behavior and lack support for complex data-driven intelligence, such as machine learning models.

Similarly, [Vale et al. 2023] introduces an MDSE framework for industrial digital twins, validating their approach through an iron ore sintering plant case study. While beneficial for industrial settings, this approach is domain-specific and does not generalize to other applications, such as agricultural or healthcare digital twins. Additionally, it is not specified whether the system supports automatic code generation capabilities.

Therefore, although significant progress has been made in relation to MDSE and digital twins, there are still a number of challenges to be addressed:

- **Lack of domain-independent digital twin frameworks:** Most existing MDSE-based digital twin proposals are designed for specific industries, limiting their adaptability and reusability across different domains.
- **Limited support for intelligent digital twin models:** Many approaches rely on predefined rule-based logic rather than incorporating machine learning and advanced simulation models. However, this integration is essential, as it enables digital twins to continuously learn from data, adapt to changing conditions, and perform intelligent tasks, thus enhancing their capabilities [Rathore et al. 2021].
- **Absence of comprehensive automation:** While some approaches enable code generation, few provide full automation, including the deployment infrastructure.
- **Fragmented digital twin service modeling:** Many frameworks do not fully define the range of services digital twins can offer, such as predictive analytics or testing of hypothetical scenarios.

Therefore, there is a clear need for comprehensive, domain-independent digital twin development frameworks. This doctoral project aims to fill this gap.

4. Research Methodology

The research methodology for this doctoral project follows a structured approach based on the principles of Design Science Research (DSR) [Hevner et al. 2010] and is organized into different phases:

- **Identify problem and motivation:** A thorough review of existing research is conducted to identify gaps and limitations in current modeling techniques for digital twins. The findings from this review provide a foundation for defining the research problem. Thus, this phase establishes the motivation for the research.
- **Define the objectives for a solution:** Based on the definition of the problem, the objectives to be met by an effective MDSE solution are established.
- **Design and development:** An MDSE approach is implemented to streamline digital twin definition, validation, generation, and deployment. For more details on the steps involved in this phase, see Section 5 where the solution is presented.
- **Demonstration:** The proposed MDSE approach is applied to selected digital twin use cases to showcase its feasibility. The purpose of this step is to illustrate how the approach can be applied in real-world situations.
- **Evaluation and refinement:** The effectiveness of the MDSE approach is systematically assessed based on predefined criteria. Insights gained from this phase can guide iterative refinements to enhance the solution.
- **Communication:** The research findings are disseminated through academic publications, presentations, and discussions within the research community. This phase ensures that the contributions of the research reach a broader audience and contribute to advancing knowledge in the field of digital twins and MDSE.

5. Proposed Solution and Preliminary Results

This section begins with an introduction to metamodeling layers. Next, the proposed solution is defined, as well as its novelty and advantages compared to existing approaches. Finally, the results and progress achieved to date are presented.

5.1. Layers of Metamodeling

In MDSE, metamodeling is structured into four hierarchical layers: M0, M1, M2, and M3 [Brambilla et al. 2017], each serving a different function (see Figure 1). The M0 layer represents actual instances of the system, such as code. The M1 layer contains models that describe the structure and behavior of M0 instances, offering an abstract representation of real-world entities. The M2 layer, or metamodel level, defines the language and rules for creating M1 models, ensuring consistency in their structure and semantics. Finally, the M3 layer establishes the fundamental rules for defining metamodels (M2).

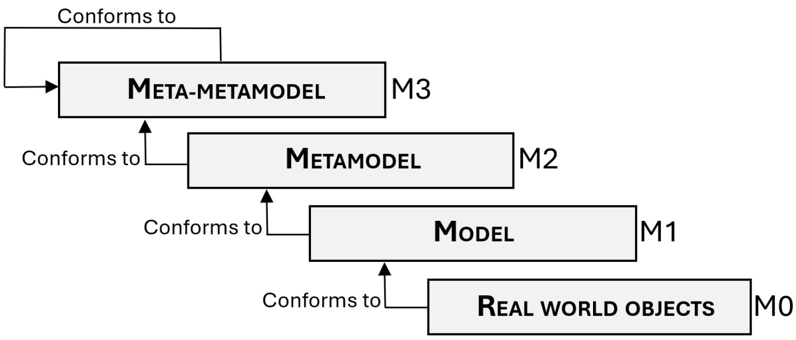

Figure 1. Layers of metamodeling. Adapted from [Brambilla et al. 2017].

5.2. Proposed Solution

Thus, to implement an MDSE approach for digital twin development, it is crucial to focus on these layers. This involves addressing the following steps:

- **Identification of the conceptual model of digital twins:** This step consists of defining in a domain-agnostic manner the fundamental characteristics and relationships of the various digital twin systems documented in the literature.
- **Metamodel (M2) formulation:** The next step involves formalizing the structure and semantics of the conceptual model of digital twins through the creation of a metamodel (M2). This metamodel provides a standardized framework that ensures consistency in the definition of digital twin models (M1). To achieve this, the Eclipse Modeling Tools [Eclipse Foundation b] are utilized.
- **Implementation of a graphical syntax for model representation:** A graphical syntax is a visual language. In this MDSE approach, this syntax allows users to design digital twin models (M1) adhering to the metamodel (M2) specifications without the need for deep technical knowledge. The Sirius [Eclipse Foundation c] and Eugenia [Kolovos et al. 2017] tools are used for this purpose.
- **M2T transformations:** Finally, M2T transformations are implemented to enable the automatic generation of all system code and deployment infrastructure (M0) from the defined digital twin models (M1). The Acceleo tool [Eclipse Foundation a] is used for this purpose.

5.3. Novelty and Advantages of the Proposed Approach

The proposed approach overcomes the limitations of related studies (see Section 3). First, it provides a domain-independent development framework for digital twins, making it applicable to a wide range of industries. Furthermore, by integrating machine learning, the proposed approach enables the creation of intelligent digital twins that go beyond predefined rule-based logic. In addition, its code generation capabilities enable the creation of all system code and deployment infrastructure, providing an end-to-end solution.

5.4. Progress and Achievements to Date

The author is in the second year of the doctoral program and has made significant progress, focusing on familiarizing with the necessary technologies and tools. A comprehensive literature review on MDSE and digital twins has also been conducted.

As part of this ongoing research, the author has published a paper in a high-impact journal [Barriga et al. 2024], where an MDSE approach to streamline the development of distributed intelligent systems is introduced. The applicability of the proposal is illustrated with an agricultural digital twin integrating various intelligent capabilities.

Moreover, in 2024, the author participated in the *Sociedad de Ingeniería de Software y Tecnologías de Desarrollo de Software* (SISTEDES) congress, where he presented in the Software Engineering for Digital Twins (ISGD) track. This experience provided valuable feedback from experts in the field.

6. Major Challenges and Future Work

One of the main challenges facing this doctoral project is the lack of universal agreement on what constitutes a digital twin, as different authors conceptualize them differently, although some elements consistently appear. This complicates the formulation of a domain-independent approach. In addition, the complexity of the technologies involved, such as IoT, cloud computing, and machine learning, presents a significant challenge.

Future work will focus on extending the capabilities of digital twins supported by the MDSE solution, including advanced prescriptive and autonomous capabilities. In this sense, prescriptive capabilities enable digital twins to recommend specific actions, while autonomous capabilities allow digital twins to execute them without human intervention. Given its complexity, the initial version overlooks these two functions.

Acknowledgements. This work was funded by project TED2021-129194B-I00 funded by MCIN/ AEI/ 10.13039/501100011033 and for European Union NextGenerationEU/ PRTR; the 0100_TID4AGRO_4_E project, co-financed by the European Regional Development Fund (ERDF) through the European Commission's INTERREG VI-A Spain-Portugal Program (POCTEP) 2021-2027; the Government of Extremadura, Council for Economy, Science and Digital Agenda under the grant GR21133; Arturo Barriga was supported by a pre-doctoral fellowship (593911006022023000186) (PD23034) awarded by the Government of Extremadura and co-financed by the European Social Fund Plus (ESF+).

References

Barriga, A., Barriga, J. A., Pérez-Toledano, M. A., and Clemente, P. J. (2024). Model-driven development towards distributed intelligent systems. *ACM Transactions on Internet Technology*.

Barriga, J. A., Chaves-González, J. M., Barriga, A., Alonso, P., and Clemente, P. J. (2023). Simulate iot towards the cloud-to-thing continuum paradigm for task scheduling assessments. *J. Object Technol.*, 22(1):1–31.

Barriga, J. A., Clemente, P. J., Sosa-Sánchez, E., and Prieto, Á. E. (2021). Simulateiot: Domain specific language to design, code generation and execute iot simulation environments. *IEEE Access*, 9:92531–92552.

Bibow, P., Dalibor, M., Hopmann, C., Mainz, B., Rumpe, B., Schmalzing, D., Schmitz, M., and Wortmann, A. (2020). Model-driven development of a digital twin for injection molding. In *International Conference on Advanced Information Systems Engineering*, pages 85–100. Springer.

Brambilla, M., Cabot, J., and Wimmer, M. (2017). *Model-driven software engineering in practice*. Morgan & Claypool Publishers.

Ciccozzi, F. and Spalazzese, R. (2016). Mde4iot: supporting the internet of things with model-driven engineering. In *International Symposium on Intelligent and Distributed Computing*, pages 67–76. Springer.

Dalibor, M., Michael, J., Rumpe, B., Varga, S., and Wortmann, A. (2020). Towards a model-driven architecture for interactive digital twin cockpits. In *International Conference on Conceptual Modeling*, pages 377–387. Springer.

Eclipse Foundation. Acceleo - generate anything from any emf model. `https://eclipse.dev/acceleo/`. Accessed: 2025-02-18.

Eclipse Foundation. Eclipse modeling tools. `https://www.eclipse.org/downloads/packages/release/2021-03/r/eclipse-modeling-tools`. Accessed: 2025-02-18.

Eclipse Foundation. Sirius - the easiest way to get your own modeling tool. `https://eclipse.dev/sirius/`. Accessed: 2025-02-18.

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

Fuller, A., Fan, Z., Day, C., and Barlow, C. (2020). Digital twin: Enabling technologies, challenges and open research. *IEEE access*, 8:108952–108971.

Hailpern, B. and Tarr, P. (2006). Model-driven development: The good, the bad, and the ugly. *IBM systems journal*, 45(3):451–461.

Hevner, A., Chatterjee, S., Hevner, A., and Chatterjee, S. (2010). Design science research in information systems. *Design research in information systems: theory and practice*, pages 9–22.

Ihirwe, F., Di Ruscio, D., Mazzini, S., Pierini, P., and Pierantonio, A. (2020). Low-code engineering for internet of things: a state of research. In *Proceedings of the*

23rd ACM/IEEE international conference on model driven engineering languages and systems: companion proceedings, pages 1–8.

Kolovos, D. S., García-Domínguez, A., Rose, L. M., and Paige, R. F. (2017). Eugenia: towards disciplined and automated development of gmf-based graphical model editors. *Software & Systems Modeling*, 16:229–255.

Lehner, D., Pfeiffer, J., Tinsel, E.-F., Strljic, M. M., Sint, S., Vierhauser, M., Wortmann, A., and Wimmer, M. (2021). Digital twin platforms: requirements, capabilities, and future prospects. *IEEE Software*, 39(2):53–61.

Macías, A., Navarro, E., Cuesta, C. E., and Zdun, U. (2023). Architecting digital twins using a domain-driven design-based approach. In *2023 IEEE 20th International Conference on Software Architecture (ICSA)*, pages 153–163. IEEE.

Mihai, S., Yaqoob, M., Hung, D. V., Davis, W., Towakel, P., Raza, M., Karamanoglu, M., Barn, B., Shetve, D., Prasad, R. V., et al. (2022). Digital twins: A survey on enabling technologies, challenges, trends and future prospects. *IEEE Communications Surveys & Tutorials*, 24(4):2255–2291.

Minerva, R., Lee, G. M., and Crespi, N. (2020). Digital twin in the iot context: A survey on technical features, scenarios, and architectural models. *Proceedings of the IEEE*, 108(10):1785–1824.

Moin, A., Rössler, S., and Günnemann, S. (2020). Thingml+ augmenting model-driven software engineering for the internet of things with machine learning. *arXiv preprint arXiv:2009.10633*.

Rathore, M. M., Shah, S. A., Shukla, D., Bentafat, E., and Bakiras, S. (2021). The role of ai, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities. *IEEE Access*, 9:32030–32052.

Segovia, M. and Garcia-Alfaro, J. (2022). Design, modeling and implementation of digital twins. *Sensors*, 22(14):5396.

Sharma, A., Kosasih, E., Zhang, J., Brintrup, A., and Calinescu, A. (2022). Digital twins: State of the art theory and practice, challenges, and open research questions. *Journal of Industrial Information Integration*, 30:100383.

Stadtmann, F., Rasheed, A., Kvamsdal, T., Johannessen, K. A., San, O., Kölle, K., Tande, J. O., Barstad, I., Benhamou, A., Brathaug, T., et al. (2023). Digital twins in wind energy: Emerging technologies and industry-informed future directions. *IEEE Access*.

Stahl, T., Völter, M., and Czarnecki, K. (2006). *Model-driven software development: technology, engineering, management*. John Wiley & Sons, Inc.

Vale, S., Reddy, S., Subramanian, S., Chaudhuri, S. R., Nistala, S. H., Deodhar, A., and Runkana, V. (2023). A model-driven approach for knowledge-based engineering of industrial digital twins. In *2023 ACM/IEEE 26th International Conference on Model Driven Engineering Languages and Systems (MODELS)*, pages 13–23. IEEE.

Wang, H., Chen, X., Jia, F., and Cheng, X. (2023). Digital twin-supported smart city: Status, challenges and future research directions. *Expert Systems with Applications*, 217:119531.